Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Structural Parameters of Scientific Knowledge Production: Panel Evidence on Homogeneity, Quality, and Global Redistribution

Johannes W. Fedderke

jwf15@psu.edu

Pennsylvania State University

Research Article

Keywords: Knowledge Production Function, endogenous growth, semi-endogenous growth, parameter homogeneity, cross-sectional dependence, AMG, CCEMG, scientific publications, human capital, citation quality, global knowledge redistribution science competition

Posted Date: July 31st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10425260/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

The Structural Parameters of Scientific Knowledge Production: Panel Evidence on Homogeneity, Quality, and Global Redistribution

Johannes W. Fedderke*

July 24, 2026

Abstract

Will the countries that lead global science today still lead mid-century? And will quality leadership prove more or less durable than volume leadership? These questions have become urgent as a wave of emerging economies has expanded their research output dramatically, while established G8 science powers face slower growth in the inputs that drive science production.

We estimate Knowledge Production Functions for 181 countries over 2003–2022, using heterogeneous-panel estimators that account for cross-sectional dependence from the global science frontier. Three structural findings emerge.

First, knowledge-stock compounding is the dominant engine of academic science output: $\hat{\phi} \in [0.65, 0.76]$ across estimators, placing the panel firmly in the semi-endogenous growth regime and substantially above patent-based estimates. We confirm long-run coefficient homogeneity across countries, so that cross-country differences in knowledge production shares reflect input accumulation histories, not heterogeneous production technologies.

Second, volume and quality of scientific output follow structurally distinct production processes. The knowledge-stock elasticity for citation impact is two orders of magnitude below its quantity counterpart ($\hat{\phi}_q \approx 0.015$ vs $\hat{\phi} \approx 0.65$): knowledge-stock compounding accrues to volume but not to citation quality. Human capital quality is the principal driver of citation impact ($\hat{\sigma}_q = 2.19$). A complementary decomposition finds that time-invariant institutional characteristics explain 65% of cross-country citation-quality variation; the near-exact inverse of the quantity specification, where measured inputs dominate.

Third, conditional projections to 2047 reveal dramatic G8 quantity-share compression. Quality projections are more nuanced: G8 quality attenuation is substantially arrested when driven by quality factor inputs alone, but approaches the quantity decline when volume compounding is incorporated. The gap between the two warns that volume decline may eventually transmit into quality leadership loss.

JEL Classification: O30, O31, O33, O47, C23

Keywords: Knowledge Production Function; endogenous growth; semi-endogenous growth; parameter homogeneity; cross-sectional dependence; AMG; CCEMG; scientific publications; human capital; citation quality; global knowledge redistribution science competition

*Pennsylvania State University, Lewis Katz Building, University Park, PA 16802, United States. E-mail: jwf15@psu.edu

1 Introduction

The distribution of scientific output, measured in both volume and quality metrics has been changing over the last quarter century. A stark pattern has been the relative rise of China paired with a relative decline of the USA, while a wave of emerging economies (India, Türkiye, South Korea) has been closing the gap on established European producers. This raises an important question: What structural parameters govern the conversion of research inputs into scientific knowledge, are they driven by efficiency differences or by differences in the accumulation of research inputs, and what do those parameters imply for the future distribution of global scientific output? These are not merely academic questions. The production of scientific knowledge is an important upstream source of technological innovation, productivity growth, and long-run economic prosperity, and the global distribution of that production is shifting rapidly. Developing answers to our questions is a prerequisite for evidence-based science policy, for accurate assessments of national scientific capacity, and for projecting the future distribution of global knowledge production. This paper focuses on the total *absolute* (not per-capita) volume of knowledge production, since aggregate national science capacity, rather than productivity per researcher, is what ultimately drives geopolitical and economic primacy.

Table 1 illustrates why these questions matter. Two features of the data stand out immediately. First, the global distribution of scientific output has shifted dramatically over the sample period: China's article share nearly tripled from 6.5% in 2003 to 21.8% by 2022, while the United States fell from 26.1% to 13.5% and Japan from 7.6% to 2.9%. India, Türkiye, and South Korea all gained ground, while the traditional European producers broadly held or modestly declined.[1] The world article total itself nearly tripled, from 1.48 million in 2003 to 4.47 million in 2022, so that stable shares mask substantial absolute gains. Second, and less visibly, the volume and citation-quality rankings diverge. The United Kingdom's global article share has fallen, yet its top-1% citation share held near 1.6%; Israel's article share declined while its citation share remained well above its output weight; and China's rapid volume gains have been accompanied by a much more modest improvement in citation impact (0.4% to 1.1%), consistent with a pattern in which quantitative catch-up precedes qualitative convergence. These two features, dramatic volume redistribution alongside persistent quality divergence, are precisely what the knowledge production function methodology (KPF) framework we employ is designed to explain, and they motivate both the quantity and quality specifications estimated in what follows.

The focus on scientific output rather than patents or R&D reflects a broader gap in the empirical literature. The innovation and productivity literature has concentrated

[1] We treat EU member states as separate national science systems: bibliometric outputs and research inputs are attributed nationally in the source data, funding is primarily national, and institutional incentives remain nationally organised.

Table 1: Scientific Output and Citation Impact of Fifteen Focus Economies, 2003–2022

		Global article share (%)			Top-1% citation share (%)		
ISO	Country	2003	2012	2022	2003	2012	2020†
G8 economies							
CHN	China	6.5	13.6	21.8	0.4	0.5	1.1
USA	United States	26.1	19.7	13.5	1.8	1.8	1.4
GBR	United Kingdom	6.6	5.3	4.3	1.3	1.6	1.6
DEU	Germany	6.5	5.5	3.9	0.9	1.2	1.0
ITA	Italy	3.5	3.2	2.9	0.6	0.9	1.3
FRA	France	4.7	3.9	2.4	0.9	1.0	0.9
CAN	Canada	3.4	3.1	2.5	1.2	1.4	1.2
JPN	Japan	7.6	4.8	2.9	0.5	0.5	0.4
Emerging market economies							
IND	India	2.0	3.3	5.3	0.5	0.5	0.7
BRA	Brazil	1.4	2.1	1.9	0.4	0.3	0.4
TUR	Türkiye	1.0	1.2	1.4	0.3	0.3	0.7
ZAF	South Africa	0.4	0.5	0.6	0.4	0.6	0.8
ISR	Israel	0.9	0.6	0.5	1.0	0.9	0.8
KOR	South Korea	1.9	2.5	2.1	0.5	0.6	0.7
MEX	Mexico	0.6	0.6	0.6	0.3	0.4	0.4
World totals (millions of articles)							
	World	1.48	2.64	4.47	–	–	–

Sources: Article counts and global shares from NSB Science and Engineering Indicators 2023 (National Science Board, 2023), whole-count basis. Top-1% citation share = each country's share of globally top-cited articles as a percentage of world top-1% output; source: NSB S&E Indicators 2023, field-normalised. World totals rounded to two decimal places.

† Citation data end in 2020; citation counts for 2021–2022 articles are still accumulating and are excluded to avoid truncation bias. This is consistent with the citation baseline used throughout the paper (see Section 5.4.2).

Note: G8 economies above the rule; emerging market economies below. Global article share computed against whole-count world total in each year.

almost exclusively on the *proximate* technological drivers of innovation (patents, R&D expenditure, and multifactor productivity) while largely neglecting the more fundamental upstream source: scientific discovery itself. Yet the economic history literature is unambiguous that it was basic scientific knowledge, embodied in the trained scientists, theoretical advances, and published findings of the seventeenth through twentieth centuries, that generated the conditions for the Industrial Revolution and subsequent productivity transformations (Mokyr, 2002; Nelson & Wright, 1992). An empirical literature on the production of scientific knowledge is therefore a missing complement to the patent- and R&D-based literature, not merely a parallel strand: it addresses the causal origin of the innovations that the commercial literature takes as its starting point.

The theoretical foundation for this inquiry is the Knowledge Production Function (KPF), originating in the endogenous growth framework of Romer (1990) and further developed in the semi-endogenous reformulation of Jones (1995a). In Romer's model, a dedicated research sector employs human capital to augment the accumulated knowledge stock, generating new knowledge. While the Romer (1990) specification imposed constant returns to both human capital and knowledge stocks, Jones (1995b) introduced the possibility of diminishing returns to both factors, resolving the potential empirical implausibility of the Romer scale effect. The key structural parameters of the KPF, the elasticity of knowledge output with respect to human capital (σ) and with respect to the accumulated knowledge base (ϕ), are jointly sufficient to discriminate among competing growth regimes. Values of $\sigma = \phi = 1$ recover the Romer specification; $\sigma, \phi \in (0, 1)$ support Jones semi-endogenous growth; and $\sigma, \phi > 1$ constitute a *strong Schumpeterian* result (Madsen, 2008), implying super-proportional returns and self-reinforcing cumulative advantage in knowledge production.

This paper provides new cross-country evidence on these parameters using an unusually rich panel dataset spanning 181 countries and two decades (2003–2022). Our principal innovation relative to the existing literature is fourfold. First, we use scientific publication counts (both whole-count and fractional-count article measures from the US National Science Board) as our primary output variable, supplemented by a logit-transformed top-1% citation share for quality analysis. This bibliometric approach complements the technology- and patent-based estimates that dominate the literature and is particularly suited to capturing the output of the academic science system, which is one proximate source of the externalities identified in the endogenous growth literature. Second, we implement heterogeneous-panel estimators, specifically the AMG estimator of Bond and Eberhardt (2009) and the CCEMG estimator of Pesaran (2006), that explicitly account for the cross-sectional dependence generated by a common global science frontier and allow for heterogeneous factor loadings across countries. The Pooled Mean Group (PMG) estimator of Pesaran et al. (1999) additionally provides a formal test of long-run coefficient homogeneity. Standard two-way fixed-effects estimators, which dominate

prior cross-country KPF work, are inconsistent under these conditions (Eberhardt & Teal, 2011; Pesaran, 2006). Third, we establish that the dominant finding from country-specific estimation is *parameter homogeneity*: the long-run KPF elasticities are statistically indistinguishable across countries, as confirmed by the PMG Hausman test. This homogeneity result validates the use of panel-mean elasticities for our forward projections and is itself an empirical contribution: it tells us that cross-country differences in knowledge production shares are driven primarily by input accumulation rates and initial conditions, not by heterogeneous production technologies.

Fourth, we identify a methodological novelty in the quality specification. The logit-transformed top-1% citation share is I(0) while the regressors are I(1), thus creating an integration-order mismatch: standard estimators, specifically FE, AMG, CCEMG, and standard PMG, are not designed for this integration-order mismatch and produce unstable estimates in practice. The estimator better suited to this setting, as it absorbs cross-sectional dependence through mean augmentation without strictly requiring a cointegrating structure between I(0) and I(1) variables, is PMG+CSD (PMG augmented with cross-sectional mean correction), which recovers a precisely estimated and economically large human capital elasticity ($\hat{\sigma}_q = 2.19$, $t = 11.2$) that the other estimators obscure. This identification strategy is, to our knowledge, new to the KPF literature.

Our main empirical findings are as follows. *First*, the knowledge stock elasticity $\hat{\phi}$ is positive, precisely estimated, and at the upper end of prior estimates: $\hat{\phi} \in [0.65, 0.76]$ across estimators in the preferred above-median sample, substantially above the patent-based estimates of Jones (1995a) ($\hat{\phi} \approx 0.3$–0.5) and consistent with absorptive capacity complementarities (Cohen & Levinthal, 1990). *Second*, human capital quality elasticity is substantial ($\hat{\sigma} \approx 2.30$ above-median; $\hat{\sigma} \approx 4.2$ in the non-zero sample), consistent with a two-stage model in which human capital drives entry into science while knowledge compounding dominates thereafter. *Third, and most distinctively*, citation-impact quality is governed by a structurally distinct production process: accumulated knowledge stocks are entirely irrelevant to citation impact ($\hat{\phi}_q \approx 0.015$, two orders of magnitude below the quantity elasticity), while human capital quality is the principal driver of citation impact ($\hat{\sigma}_q = 2.19$, $t = 11.2$) under PMG+CSD, the cross-sectional-dependence-corrected PMG estimator that is the methodologically more appropriate choice here. The citation-quality dependent variable is I(0) while the regressors are I(1), an integration-order mismatch that renders FE, AMG, CCEMG, and standard PMG problematic for this specification, as Section 5.3 documents; PMG+CSD's cross-sectional mean augmentation addresses the mismatch directly without requiring a cointegrating structure. A complementary country fixed-effects decomposition shows that a time-invariant institutional characteristic also explains 65% of the cross-country *level* of citation quality beyond measured inputs, a near-complete inversion of the quantity specification, where slopes explain 63% of variance and the fixed effect only 5%. Quality leadership is therefore both more contestable

than quantity leadership (no compounding protection) and shaped jointly by human capital quality and a distinct institutional component governing how effectively that human capital translates into citation impact. *Fourth*, the dominant finding from the country-specific coefficient analysis is parameter homogeneity: the cross-country distribution of KPF elasticities is tight enough that PMG pooling is not rejected, and most pairwise country differences are not robustly distinguishable at $T = 20$.[2]

Forward projections to 2047 under current input trajectories flag a dramatic redistribution of scientific output: the US volume share falls from 14% to under 3%, while China rises toward 28%. Quality projections distinguish two approaches. Projecting top-1% citation shares directly from quality KPF factor inputs (Approach A, primary), G8 quality attenuation is substantially arrested: the US retains 17.1% of world top-cited science by 2047, declining only -4.4 pp vs -11.0 pp in volume. However, when volume dynamics are additionally incorporated (Approach B, qualifier), the quality-weighted decline approaches the volume decline (-16.4 pp for the US), reflecting the compounding mechanism operating through the volume channel. The gap between the two approaches is a warning against complacency: volume decline will eventually crack quality leadership through channels not modelled in the quality KPF.

A note on the forward projections. The 25-year projection exercises in Section 5 use panel-mean KPF elasticities and country-specific input accumulation trends from the 2003–2022 sample. They are best understood as conditional scenario projections (rather than forecasts): detailing the quantitative implications of carrying current trajectories forward under structural parameter stability, not as unconditional point predictions. Their purpose is to discipline the policy discussion: every historical displacement of a scientific leader was preceded by decades during which the incumbent underestimated the pace of catch-up, and the projections quantify precisely what "policy as usual" implies under the estimated KPF structure.

The remainder of the paper is organised as follows. Section 2 surveys the theoretical and empirical background. Section 3 presents the empirical specifications and variable construction. Section 4 describes the estimation methodology. Section 5 reports the main estimation results and forward projections. Section 6 concludes with policy implications. Appendix A presents a fuller treatment of the background literature surrounding scientific leadership and its displacement. Appendix B presents the full set of country-specific coefficient scatter plots for all fifteen focus economies (eight G8 and seven emerging market economies), together with the econometric motivation for the AMG and CCEMG estimators (B.1) and Monte Carlo evidence on their performance (B.2). Appendix C reports country-specific vs panel-mean projection comparisons. Appendix D discusses the methodological implications of using rest-of-world knowledge stocks for the estimators employed. Appendix E reports the full complement of panel cointegration test statistics.

[2]We do discuss some country-specific insights where appropriate.

2 Endogenous Growth, Knowledge Production Functions, and Human Capital

The literature on endogenous growth and KPFs is substantial, and our discussion here is intended to do no more than highlight the features to emerge that are of immediate significance to the analysis presented by the paper. A more extended survey is presented in the Online Appendix A.

Endogenous Growth: the Knowledge Production Function A key feature of Romer (1986) is the treatment of knowledge as a non-rival, non-excludable public good, generating aggregate increasing returns that sustain endogenous growth. Romer (1990) endogenised knowledge production itself, specifying a dedicated research sector in which human capital and the existing knowledge stock jointly determine the rate of new knowledge creation, the relationship formalised in equation (1) below. The Romer (1990) formulation implies a strong scale effect: a larger research workforce produces proportionally more innovation. Jones (1995b) subjected this prediction to time-series evidence and rejected it, motivating the semi-endogenous reformulation of Jones (1995a), which allows diminishing returns to both researchers ($\sigma < 1$) and the knowledge stock ($\phi < 1$).[3] An alternative resolution (Young (1998), Kortum (1997), Segerstrom (1998)) attributes the absence of scale effects to the endogenous expansion of product variety, which spreads research effort over an ever-larger number of product lines; Jones (1995a) introduces Q explicitly as the product-variety dilution term in the KPF. The quality-ladder framework of Aghion and Howitt (1992), in which knowledge advances along quality ladders so that the discovery of new knowledge comes at the cost of destroying previously established knowledge, might predict that knowledge stock compounding should matter less for the quality of scientific output than for its volume, since creative destruction partially displaces the value of accumulated knowledge, consistent with our finding that $\hat{\phi}$ dominates quantity specifications but is entirely insignificant in quality equations.[4]

The empirical implication is that KPF parameters $\{\sigma, \phi\}$ discriminate among competing frameworks: $\sigma = \phi = 1$ corresponds to Romer (1990); $\sigma, \phi \in (0, 1)$ to Jones (1995); and $\sigma, \phi > 1$ to the 'strong Schumpeterian' result of Madsen (2008). Rather than imposing either restriction, we treat ϕ as a free parameter and test both.

[3] Segerstrom (1998) derives a fully specified semi-endogenous model with a balanced growth path consistent with the Jones (1995b) time-series evidence.

[4] The link to our empirical specification is direct: to the extent that highly cited papers reflect genuinely new knowledge, and new knowledge displaces old knowledge in the quality-ladder sense, highly cited papers are correlated with the creative destruction process that Aghion and Howitt were formalising, which in turn implies that their production is less dependent on inherited knowledge stocks than raw article volume.

Empirical Evidence, Human Capital, and Knowledge Measurement We distinguish our object of interest, the *academic* KPF relating research-sector human capital and accumulated scientific knowledge to article output, from the *aggregate innovation* KPF relating R&D expenditure to commercial patenting (Hall et al., 2001; Pakes & Griliches, 1984). The distinction corresponds to Mokyr (2002)'s differentiation between λ (practical) and Ω (epistemic) knowledge. Empirical aggregate innovation KPF estimation originates with Griliches (1990) and Griliches and Mairesse (1998), with a systematic cross-country test of the Romer–Jones distinction presented in Madsen (2008), who found strong Schumpeterian estimates ($\hat{\phi} > 1$) in OECD historical panels.[5] However, still for aggregate innovation findings, Aghion et al. (2017) and Bloom et al. (2020) document declining research productivity in the US, consistent with the Jones sub-proportional regime at the frontier. By contrast to aggregate innovation results, scientific KPF evidence is harder to locate, though bibliometric measures of knowledge output are developed in Narin et al. (1976) and Larue et al. (2022); the choice between whole- and fractional-count attribution matters significantly in multi-author, internationally collaborative science (Waltman et al., 2015).

Human capital in the KPF. Human capital (HC) enters the KPF directly as the research workforce input. We use research-sector human capital (H_A), not economy-wide schooling, as the theoretically motivated input, following Romer's original specification. Benhabib and Spiegel (1994) further showed it also facilitates technology adoption from the frontier. Since Hanushek and Woessmann (2012) establish that quality-based measures have substantially larger growth effects than years-of-schooling proxies,[6] we also employ a quality-adjusted human capital index alongside raw researcher counts.

Absorptive capacity. The hypothesis that the *absorptive capacity* of economies implies that countries with larger knowledge stocks are better placed to exploit globally available knowledge has an extended history (Cohen and Levinthal (1990), Gerschenkron (1962), Abramovitz (1986)). Absorptive capacity provides a natural explanation for a two-stage model of scientific knowledge production: at the extensive margin, entry of low-income countries into science, human capital quality is the binding constraint, generating large returns to the first increments of HC quality; among established science systems, accumulated knowledge raises researcher productivity through the $\hat{\phi}$ channel and becomes the primary structural determinant of output growth. The two-stage logic predicts that HC elasticities should be larger at the extensive margin than among established science systems. Talent migration amplifies these returns by expanding the receiving country's

[5] Fedderke and Liu (2017) replicate the Madsen patterns for South Africa, motivating our heterogeneous-panel approach.

[6] Pritchett (2001) documents that raw schooling expansions often fail to translate into productivity gains, reinforcing the case for quality-adjusted HC.

effective human capital stock.

Knowledge Spillovers, the Global Frontier, and Cross-Sectional Dependence Insofar as scientific knowledge has any public-good character, a common global science frontier generates cross-sectional dependence in panel errors. Coe and Helpman (1995) documented significant R&D spillovers across bilateral trading partners; Coe et al. (2009) confirmed and extended these results; and Leydesdorff and Wagner (2009) showed that a small core of leading countries produces a disproportionate share of internationally cited science, consistent with knowledge spillovers flowing outward from the frontier. The econometric implication is critical: standard fixed-effects estimation is inconsistent when unobserved common factors (the global science frontier) have heterogeneous factor loadings (Eberhardt & Teal, 2011; Pesaran, 2006). Our preferred AMG and CCEMG estimators explicitly absorb this common frontier process: AMG via a non-parametric Common Dynamic Process, CCEMG via cross-sectional mean augmenters. Moreover, AMG and CCEMG are consistent under cross-sectional dependence that standard estimators cannot handle. The full econometric argument and Monte Carlo evidence are in Online Appendix B.

National Innovation Systems, Science Leadership, and Historical Geography The KPF's country fixed effect μ_i encodes national innovation system (NIS) quality (Nelson, 1993): the institutional framework (universities, funding agencies, IP regimes, industry-science linkages) that determines how efficiently inputs are converted into outputs. Section 5.2 interprets $\hat{\mu}_i$ as a cross-country NIS ranking.

The long-run implications are shaped by catch-up dynamics. Gerschenkron (1962)'s 'advantage of backwardness' posits that laggard countries can grow faster than the frontier by adopting developed technologies; Fagerberg (1994) extended this to the innovation context, though catch-up is conditional on absorptive capacity (Cohen & Levinthal, 1990): the potential is realised only when the lagging country has sufficient capacity to implement frontier technologies.

The historical expression of these dynamics is the sequential succession of scientific leaders.[7] The record of the past four centuries traces a recognisable pattern: English natural philosophy in the seventeenth century, French institutional science in the eighteenth, German research universities in the nineteenth, and American dominance from the mid-twentieth century onward (Ben-David, 1971). Ben-David attributes the displacement of scientific leaders to the ability of the national science system to solve three interrelated problems: the establishment of a distinct professional role for the scientist, the achievement of intellectual autonomy for scientific inquiry, and the creation of competitive

[7]The economic history treatment of scientific leadership and production is particularly extensive and rich. The discussion here is painfully sparse. A much fuller treatment is presented in Online Appendix A.

incentives for original discovery.[8]

The displacement of German scientific leadership by the United States is the transition providing the most recent evidence (Heinze et al. (2019), von der Heyden et al. (2025)). Nelson and Wright (1992) locate the transition in the maturation of the American research university and the disruption of German science under National Socialism; Moser et al. (2014) provide causal evidence that Jewish émigrés from Nazi Germany increased US patenting by 31 per cent in their fields, requiring the existing American knowledge base to be sufficiently deep to absorb the incoming human capital, likely a lower bound estimate of the aggregate impact given the substantial proportion of émigrés who were not Jewish. Borjas and Doran (2012) document a parallel dynamic from the post-1991 outflow of Soviet mathematicians, finding increases in aggregate mathematical output alongside competitive displacement of some incumbent American researchers. Mokyr (2002) connects these patterns to the core mechanism: the knowledge stock embodied in trained scientists and research institutions raises the marginal productivity of contemporaneous research, and no country has achieved sustained scientific leadership without first accumulating such a stock — the historical analogue of the $\hat{\phi}$ compounding we estimate.

Three implications follow for the present analysis. First, $\hat{\phi}$ in equation (1) is the central parameter for long-run leadership dynamics. Second, the historical quantity-before-quality sequence implies China's current pattern is structural rather than anomalous. Third, talent migration is the historically fastest displacement mechanism.

China's Scientific Ascent: Facts, Drivers, and Quality China became the world's largest single-country producer of scientific articles by 2016 (Tollefson, 2018), approaching parity with the United States on a fractional-count basis. The quality dimension has attracted scrutiny: Hu and Jefferson (2009) found that China's patent explosion was partly driven by policy incentives rather than genuine inventive activity; Yu et al. (2022) showed that citation-impact gaps between Chinese and US publications remain large and have narrowed only slowly; and Meng and Romanowich (2021) document convergence in quantitative output faster than in quality-weighted indicators. Fu (2015) provides the institutional account, arguing that China's scientific rise reflects deliberate state deployment through targeted funding programmes (the 985 and 211 Project universities, Thousand Talents, and systematic targeting of journal publications in performance evaluation). If

[8]Several analytical mechanisms have been advanced to explain these historical transitions. The economic history tradition (Abramovitz, 1986; Gerschenkron, 1962) identifies the advantage of backwardness: laggard countries that have accumulated sufficient social capabilities can grow faster than the frontier by adopting technologies developed elsewhere. Implicitly Nelson and Wright (1992) adopt this for the scientific context. The sociology of science tradition (Ben-David, 1971; Hollingsworth & Hollingsworth, 2011) emphasises that scientific creativity flourishes in decentralised, competitive institutional environments, and that once a science system achieves dominance it tends to become conservative. Mokyr (2002) connects to the deeper mechanism of knowledge accumulation: the 'useful knowledge' stock raises the marginal productivity of contemporaneous research, and no country has achieved sustained scientific leadership without first accumulating such a stock.

this expansion reflects primarily an increase in input quantity rather than structural efficiency, China's KPF coefficients should not differ significantly from the panel average, which is indeed our finding. A related concern is that explicit per-publication financial incentives in Chinese universities (Quan et al., 2017) may have inflated measured quantity and distorted citation behaviour, so that some of China's output expansion reflects incentive responses rather than genuine knowledge production; both the quantity and quality signals for China should be read with this caveat in mind.

International knowledge spillovers (Coe & Helpman, 1995) and the possibility of directed technical change (Acemoglu (2002); as China's research workforce grows, global research effort may shift toward areas of Chinese comparative advantage) introduce caveats about interpreting aggregate KPF coefficients over a twenty-year window. Our estimates establish an empirical baseline for 2003–2022 without ruling out differential dynamics over a longer horizon.

3 Empirical Specifications

3.1 Theoretical Framework

Our empirical specification is grounded in the endogenous growth literature initiated by Romer (1990) and subsequently refined by Jones (1995a). The quality-ladder (Aghion & Howitt, 1992) and product-variety (Jones, 1995a; Kortum, 1997; Segerstrom, 1998; Young, 1998) alternatives share the same fundamental KPF structure, so the specification below nests all three traditions. Knowledge accumulates via a knowledge production function (KPF):

$$\dot{A}_{it} = \delta_{it} \left(\frac{H_{A,it}}{Q_{it}} \right)^{\sigma} A_{it}^{\phi}, \qquad 0 \leq \sigma, \quad 0 \leq \phi, \tag{1}$$

where $\dot{A}_{it}$ is the time-rate of change of the knowledge stock for country i in period t; $H_{A,it}$ is human capital employed in research; Q_{it} is a measure of product variety, following Jones (1995a); A_{it} is the accumulated knowledge stock; and $\delta_{it} > 0$ is a country- and time-specific research productivity factor, capturing heterogeneity in national innovation systems and secular changes in research efficiency (Romer, 1990) (absorbed into the country μ_i and time λ_t fixed effects upon log-linearisation, so δ_{it} does not appear in the estimating equations below). The product variety term Q_{it} captures the observation of Jones that the expansion of the number of product lines over which knowledge must be generated dilutes the effective impact of the research workforce $H_{A,it}$ on aggregate innovation, thereby preventing the unbounded growth that strict Romer-type proportionality between research effort and innovation would otherwise imply.[9]

[9] A natural academic-science equivalent of product variety might be the number of active Web of Science subject categories per country-year, which would proxy for the breadth of research fields over which effort is distributed. This measure is not available in the current cross-country panel; constructing

The parameter $\sigma \geq 0$ governs the elasticity of knowledge production with respect to research inputs and is central to the distinction between the fully endogenous Romer specification ($\sigma = 1$) and the semi-endogenous formulation of Jones ($0 \leq \sigma < 1$), with $\sigma > 1$ constituting a strong-Schumpeterian regime for human capital. The parameter ϕ captures the elasticity of knowledge production with respect to the accumulated knowledge stock. The Romer specification imposes $\phi = 1$ (proportional returns); Jones semi-endogenous growth requires $0 \leq \phi < 1$ (sub-proportional); $\phi > 1$ constitutes a 'strong Schumpeterian' result in the terminology of Madsen (2008), consistent with the findings of Fedderke and Liu (2017). A symmetrical strong-Schumpeterian extension applies to the human capital input, allowing $\sigma > 1$ (super-proportional returns to research workforce quality). Rather than imposing either the Romer or Jones restrictions, the present paper treats both ϕ and σ as free parameters to be estimated, with $\phi \lessgtr 1$ and $\sigma \lessgtr 1$ as testable hypotheses.

3.2 Empirical Specification

Taking logarithms of equation (1), our baseline empirical specification is:

$$\ln \dot{A}_{it} = \mu_i + \lambda_t + \sigma \ln H_{A,it} - \sigma \ln Q_{it} + \phi \ln A_{i,t-1} + \varepsilon_{it}, \tag{2}$$

where ϕ is the knowledge stock elasticity (the key parameter whose position relative to unity discriminates among the Romer, Jones, and strong-Schumpeterian regimes); μ_i denotes country fixed effects capturing time-invariant determinants of research productivity (institutional quality, geographic factors, cultural norms; note also that μ_i absorbs the research productivity factor δ of equation (1) upon log-linearisation); and λ_t denotes year fixed effects absorbing common global shocks to scientific output (the global publication norm, changes in academic incentive structures, advances in communication technology). The error term ε_{it} is assumed to satisfy standard regularity conditions.

Three features of the specification warrant explicit comment. First, we use the *lagged* knowledge stock $A_{i,t-1}$ rather than the contemporaneous level. This is necessary to avoid a mechanical positive correlation: our measure of the knowledge stock is itself constructed as a perpetual-inventory accumulation of the dependent variable, so using the contemporaneous stock would inflate the estimated ϕ by construction.

Second, the dependent variable $\ln \dot{A}_{it}$ is operationalised as the natural logarithm of the annual count of scientific articles produced by country i in year t. In the Romer–Jones framework, the flow of new knowledge is the innovation output of the knowledge production sector, and annual article output is the closest observable proxy available in the panel data. This is consistent with the approach of Guellec and van Pottelsberghe de la Potterie (2001) and Ülkü (2004), who use patent counts, and with the bibliometric KPF literature surveyed in Hall et al. (2010).

it would require field-level bibliometric data disaggregated by country, representing a natural extension.

Third, equation (2) implies a sign restriction: since σ enters with opposing signs on H_A and Q, the estimated elasticities with respect to $\ln H_{A,it}$ and $\ln Q_{it}$ should be equal in magnitude and opposite in sign. The data reject this restriction, and we therefore treat both coefficients as free parameters.[10]

3.3 Sample Selection and Variable Construction

Dependent variables. We employ three alternative measures of knowledge output:

1. *Whole-count articles* (W_{it}): the count of scientific and engineering articles attributed in full to every contributing country, as reported in the US National Science Board Science and Engineering Indicators 2023 (National Science Board, 2023).

2. *Fractional-count articles* (F_{it}): the same publication database, but with each article distributed proportionally across contributing countries according to institutional affiliation. Fractional counting avoids the double-counting of internationally co-authored work and provides a more conservative measure of country-attributed output (National Science Board, 2023).

3. *Logit-transformed top-1% citation share* (T_{it}): the share of a country's publications appearing in the top percentile of citation-weighted articles, as reported in NSB Table SPBS-54 (National Science Board, 2023).[11] This variable is expressed as a proportion and is transformed via the logit mapping $T^*_{it} = \ln[T_{it}/(1 - T_{it})]$ prior to estimation. It serves as the dependent variable in the *quality specification*, intended to capture the effectiveness of the knowledge production process in generating high-impact research, as distinct from raw volume. A percentile share is preferred over a mean normalised citation score because citation distributions are extremely right-skewed (the mean is dominated by a small number of outlier papers and is unstable across fields and time) while the top-percentile share is a robust statistic of precisely the part of the distribution that is of interest (Waltman et al., 2012).

The primary quantity-based analysis crosses these two output measures with two human capital proxies (see Human Capital Inputs below), generating four quantity specifications: Specifications A and B use whole-count articles (W_{it}) as the dependent variable (A with raw researcher count, B with PWT human capital index); Specifications C and D use fractional-count articles (F_{it}) with the same human capital pairings. A fifth specification

[10] A Wald test of H_0: $\hat{\sigma} + \hat{\gamma} = 0$ is rejected in AMG Spec B at the 1% level ($\hat{\sigma} + \hat{\gamma} = 2.25$, SE ≈ 0.81, $t = 2.8$) and in 2 of 6 estimator–specification combinations at 5%. The sign instability of $\hat{\gamma}$ documented in Section 5.2 provides further independent evidence that the restriction does not hold in the data.

[11] Note that the NSB bibliometric data are sourced from Web of Science, which systematically undercovers non-English journals, regional publications, and the social sciences and humanities (Mongeon & Paul-Hus, 2016); this coverage bias may understate the true science output of India, Brazil, Türkiye, and South Africa relative to established Anglophone systems, and future work drawing on broader databases such as OpenAlex would provide a useful robustness check.

uses the logit citation share (T^*_{it}) as the dependent variable and is presented as the quality specification, estimated separately on the above-median sample (see Sample Construction below). For the quality specification we use the PWT human capital index as the preferred human capital proxy (Spec Q-B), and also report a researcher-count variant (Spec Q-A) for completeness. The choice of the PWT index for the quality specification is driven by two considerations: its broader country coverage (enabling a larger above-median panel) and its quality-adjusted construction, which is conceptually more coherent with a citation-impact outcome than a raw count of researchers.

Human capital inputs (H_A). We employ two alternative proxies for research-sector human capital:

1. *Raw researcher count*: derived by multiplying the World Bank World Development Indicators series 'Researchers in R&D (per million people)' (World Bank, 2026) by total population expressed in millions (i.e., total population in persons divided by 10^6), yielding an absolute stock of full-time-equivalent researchers. The estimated human capital elasticity $\hat{\sigma}$ therefore represents the percentage change in annual article output associated with a 1% increase in the absolute number of researchers. This variable most directly maps to the Romer–Jones concept of H_A as the number of persons engaged in knowledge production.

2. *PWT human capital index*: the quality-adjusted human capital measure from Penn World Tables version 11.0 (Feenstra et al., 2015), constructed from years of schooling and returns to education following Psacharopoulos and Patrinos (2004). This variable captures the *quality* dimension of the research labour force rather than its size, and is available for a substantially broader country sample than the researcher count.

The researcher-count specification is estimated on the sample of countries with adequate researcher data coverage (Specifications A and C), while the PWT human capital index is used in the broader-sample specifications (Specifications B and D).

Product variety (Q). Following Fedderke and Liu (2017) and Guellec and van Pottelsberghe de la Potterie (2001), we proxy product variety by a perpetual-inventory stock of total patent applications by applicant origin, sourced from the WIPO IP Statistics Data Centre (World Intellectual Property Organization, 2026). [12] The patent stock is constructed as:

$$P_{i,t} = (1 - \rho_P)\, P_{i,t-1} + F^P_{i,t}, \tag{3}$$

[12] Madsen (2008) constructs the patent stock via a perpetual inventory method applied to OECD patent data, with a 15% depreciation rate: essentially the same approach as ours.

where $F^P_{i,t}$ is the annual flow of patent applications and $\rho_P = 0.15$ denotes the depreciation rate on patent stock, following Hall et al. (2005) and Millot (2009).[13] Robustness checks replace the patent stock with the patent application *flow*, real GDP, and population as alternative proxies for Q.

Knowledge stock (A). The knowledge stock is constructed by perpetual inventory ($\rho_K = 0.10$ denotes the knowledge stock depreciation rate, Coe and Helpman 1995), initialised at steady state using the article series from 1980. Robustness checks for $\rho_K \in \{0.05, 0.15\}$ are in Table 6.

Rest-of-world knowledge stock (K^{-i}). To test the global public good hypothesis described in Section 2, we additionally construct for each country the rest-of-world knowledge stock:

$$K^{-i}_{t-1} = K^{world}_{t-1} - K_{i,t-1} = \sum_{j \neq i} K_{j,t-1}, \tag{4}$$

where $K^{world}_{t-1} = \sum_j K_{j,t-1}$ is the sum across all countries in the above-median sample. The augmented specification including both own-stock and rest-of-world compounding is:

$$\ln \dot{A}_{it} = \mu_i + \lambda_t + \sigma \ln H_{A,it} + \gamma \ln Q_{it} + \phi_1 \ln K_{i,t-1} + \phi_2 \ln K^{-i}_{t-1} + \varepsilon_{it}, \tag{5}$$

where ϕ_1 captures own-stock compounding and ϕ_2 the international knowledge spillover. This specification is estimated as a robustness exercise (results discussed in Section 5.2 and reported in full in Appendix D).

Note that K^{world}_{t-1} itself cannot serve as a regressor: it is identical for all countries in a given year, so it is exactly collinear with the year fixed effects in FE and with the time-invariant component of the Common Dynamic Process in AMG and the cross-sectional means in CCEMG. In other words, whatever global trend each estimator is designed to absorb, it will absorb K^{world}_{t-1} along with it.

The rest-of-world stock K^{-i}_{t-1} avoids exact collinearity because the log transformation is strictly concave:

$$\ln K^{-i}_{t-1} = \ln\!\big(K^{world}_{t-1} - K_{i,t-1}\big) \;\neq\; \ln K^{world}_{t-1} - \ln K_{i,t-1}. \tag{6}$$

For small countries, where $K_{i,t-1} \ll K^{world}_{t-1}$, we have $\ln K^{-i}_{t-1} \approx \ln K^{world}_{t-1}$: effectively the same as the world stock and absorbed along with it. For the five or six largest science economies (United States, China, Germany, Japan, United Kingdom), which together account for roughly 50–60% of the world knowledge stock, the subtraction is material and $\ln K^{-i}_{t-1}$ diverges meaningfully from $\ln K^{world}_{t-1}$. The cross-sectional variation of

[13] The patent stock is initialised at $P_{i,t_0} = F^P_{i,t_0}/(g_i + \rho_P)$, where g_i is the mean log-growth rate of patent applications over the first five observed years of data for country i.

$\ln K_{t-1}^{-i}$ within each year is therefore small (within-year cross-sectional standard deviation of 0.036), and identification of the spillover coefficient $\hat{\phi}_2$ (the coefficient on the rest-of-world knowledge stock $\ln K_{t-1}^{-i}$ in the augmented specification) relies primarily on the gradient across large-economy outliers. The consequences for estimation precision, and for the behaviour of each estimator in the augmented specification, are discussed in Section 4.2.

Sample construction. The estimation sample spans 2003–2022, determined by the intersection of bibliometric coverage (NSB S&E Indicators, which begin in 2003) and the 2022 endpoint of the most recent release. Two nested samples are constructed:

- *Non-zero sample*: all countries reporting at least one scientific article over the full 2003–2022 period (181 countries).
- *Above-median sample*: countries whose time-averaged annual article output exceeds the cross-country median of approximately 677 articles per year (90 countries). This restricts the sample to science-active economies with sufficient bibliometric mass for stable estimation.

Left-censoring and sample selection. A potential concern is that the two-sample design may introduce selection bias. Countries that produce zero articles in some years are observable in the non-zero sample only when they produce positive output; the below-median countries are excluded from the above-median sample entirely. A Heckman (1979) selection correction would be appropriate if we were estimating the binary decision to become a science-producing country, but our object of interest is the conditional KPF: how much science is produced given that a country is already an active producer. The non-zero sample (181 countries) addresses this by including all observable producers; the above-median sample restricts to countries with sufficient bibliometric mass for stable KPF estimation, where selection into the sample is determined by the level of science activity rather than unobservables. The comparison of results across the two samples (Section 5.2) provides an informal check: if selection bias were material, we would expect systematic differences in coefficient patterns beyond those attributable to the extensive vs intensive margin distinction. The patterns we observe are interpretively consistent with the two-stage science development model rather than selection artefacts.

Table 2 presents descriptive statistics for the key estimation variables.

3.4 Estimation Specifications

We estimate four primary specifications, crossing two output measures (whole count, fractional count) with two human capital proxies (researchers, PWT hc):

Table 2: Descriptive Statistics, 2003–2022

	Non-zero sample ($N = 181$)			Above-median sample ($N = 90$)		
Variable	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Articles (whole count)	16,127	677	58,842	30,982	7,307	78,995
Articles (fractional count)	12,475	376	49,486	24,029	5,176	66,785
Top-1% citation share (%)	0.53	0.42	0.56	0.60	0.52	0.52
Researchers (FTE, thousands)	95,064	17,194	256,507	117,897	28,756	281,598
PWT human capital index	2.61	2.62	0.71	2.89	2.97	0.58
Patent stock (applications)	16.8M	2,543	218.7M	23.7M	8,390	269.2M
Knowledge stock (whole, PIM)	146,234	4,060	616,148	206,649	39,294	544,795
RoW knowledge stock (K^{-i})	25,290,264	24,166,974	3,048,524	18,386,286	17,538,700	5,937,913

Sources: NSB (National Science Board, 2023); WIPO (World Intellectual Property Organization, 2026); WDI (World Bank, 2026); PWT 11.0 (Feenstra et al., 2015). Knowledge and patent stocks computed by authors via PIM. FTE = full-time equivalent.

$$\textbf{Spec A:} \quad \ln W_{it} = \mu_i + \lambda_t + \sigma \ln R_{it} + \gamma \ln P_{i,t} + \phi \ln K^W_{i,t-1} + \varepsilon_{it} \tag{7}$$

$$\textbf{Spec B:} \quad \ln W_{it} = \mu_i + \lambda_t + \sigma \ln H_{it} + \gamma \ln P_{i,t} + \phi \ln K^W_{i,t-1} + \varepsilon_{it} \tag{8}$$

$$\textbf{Spec C:} \quad \ln F_{it} = \mu_i + \lambda_t + \sigma \ln R_{it} + \gamma \ln P_{i,t} + \phi \ln K^F_{i,t-1} + \varepsilon_{it} \tag{9}$$

$$\textbf{Spec D:} \quad \ln F_{it} = \mu_i + \lambda_t + \sigma \ln H_{it} + \gamma \ln P_{i,t} + \phi \ln K^F_{i,t-1} + \varepsilon_{it} \tag{10}$$

where W_{it} and F_{it} are whole- and fractional-count articles; R_{it} is the raw researcher count; H_{it} is the PWT human capital index; $P_{i,t}$ is the patent stock; and $K^W_{i,t-1}$ and $K^F_{i,t-1}$ are the lagged knowledge stocks constructed from the corresponding article series.

A fifth specification uses the logit-transformed top-1% citation share T^*_{it} as the dependent variable in place of W_{it} and F_{it}, and is estimated on the above-median sample only, since the citation share is uninformative for countries at the periphery of global science. Note that in the quality specification Q does not carry the Jones (1995) product-variety dilution interpretation it bears in the quantity equation: since the dependent variable is a citation *share* rather than an absolute count, the variety-spreading mechanism does not apply, and $\hat{\gamma}_q$ is better interpreted as capturing complementarity between the commercial knowledge environment and high-impact basic research. Thus instead, the density of patentable innovation in a national research environment raises the probability that academic work in that environment produces frontier-cited output, consistent with the science-technology interface documented in Narin et al. (1997).

4 Empirical Methodology

4.1 Cross-Sectional Dependence and the Inadequacy of Standard Estimators

Knowledge production is an inherently global process: a common science frontier generates cross-sectional dependence in panel errors.[14] Standard fixed effects (FE) accounts for this only under homogeneous loadings, untenable when countries occupy different frontier positions; heterogeneous loadings render FE inconsistent with severely understated standard errors (Eberhardt & Teal, 2011; Pesaran, 2006).[15]

4.2 Panel Unit Root Tests

Given the cross-sectional dependence documented above, we employ the Pesaran (2007) Cross-sectionally Augmented IPS (CIPS) test, which augments each country's ADF regression with the cross-sectional mean and its lag:

$$\Delta y_{it} = \alpha_i + b_i\, y_{i,t-1} + c_i\, \bar{y}_{t-1} + d_i\, \Delta \bar{y}_{t-1} + e_{it}, \tag{11}$$

where $\bar{y}_t = N^{-1} \sum_i y_{it}$ is the cross-sectional mean. The CIPS statistic is the simple average of the individual Cross-sectionally Augmented Dickey-Fuller (CADF) t-statistics: $\text{CIPS} = N^{-1} \sum_{i=1}^{N} t_i(b_i)$, with critical values tabulated in Pesaran (2007). The CIPS test has power advantages over Levin–Lin–Chu and Im–Pesaran–Shin tests under cross-sectional dependence (Kapetanios et al., 2011).

Table 3 presents the CIPS statistics for all variables in levels and first differences. The results confirm that the knowledge stock and the human capital index are I(1) processes, while the logit-transformed top-1% citation share is stationary in levels. Annual article output is I(1) in the above-median sample (where all countries exhibit clear upward trends) and borderline I(0) in the full non-zero sample (where the heterogeneous mix of growing and stable peripheral countries weakens the trend signal). The patent stock appears I(0) in levels, consistent with its construction as a geometric sum of flows that are bounded by country absorptive capacity.

The mixed integration orders across variables raise questions about the treatment of specifications that combine I(0) and I(1) variables. For the *quantity* specifications, in which the dependent variable ($\ln \dot{A}_{it}$, I(1) in the above-median sample) and the key re-

[14] Formally, $\varepsilon_{it} = \boldsymbol{\lambda}_i' \mathbf{f}_t + u_{it}$ where $\mathbf{f}_t$ are unobserved common factors and $\boldsymbol{\lambda}_i$ country-specific loadings.

[15] The author used Claude (Anthropic, Sonnet 4.6) during the preparation of this manuscript. The AI's role was across two dimensions. First, it assisted with the copy editing of prose. All substantive arguments, interpretations, and theoretical claims originate with the author and were subject to the author's review and revision at each stage. Second, AI was used in the checking and cleaning of code. All numerical results were verified by the author. The author remains fully responsible for the research design, the interpretation of all results, and the intellectual content of the paper in its entirety.

Table 3: CIPS Panel Unit Root Test Results

Variable	Non-zero ($N = 181$) Levels	Non-zero ($N = 181$) First diff.	Above-median ($N = 90$) Levels	Above-median ($N = 90$) First diff.
ln(Articles, whole)	-2.426^{*}	-4.724^{***}	-1.276	-3.501^{***}
ln(Articles, frac)	-2.502^{**}	-4.751^{***}	-1.588	-3.634^{***}
Logit: Top-1%	-3.413^{***}	-5.874^{***}	-3.507^{***}	-6.045^{***}
ln(Researchers)	-1.250	-3.671^{***}	-1.612	-3.952^{***}
ln(PWT hc index)	-1.729	-1.994	-1.647	-2.338^{**}
ln(Patent stock)	-3.623^{***}	-2.539^{**}	-3.308^{***}	-2.323^{*}
ln(Knowledge stock)	-0.961	-2.657^{**}	-0.368	-0.950
ln(K^{-i}, RoW)	0.014	-2.445^{**}	-4.700^{***}	-1.075
5% critical value		-2.33		
1% critical value		-2.57		

Notes: CIPS statistic (average of country-level CADF t-statistics). Null hypothesis: all panels contain a unit root. Specification includes a constant but no trend. Critical values from Pesaran (2007), Table 2. ***, **, * denote rejection at 1%, 5%, 10%.

gressors (knowledge stock, human capital) are all I(1) and plausibly form a cointegrating relationship consistent with the KPF framework (Engle & Granger, 1987), FE, AMG, CCEMG, and PMG are all appropriate, subject to the qualifications discussed in Section 4.3. Panel cointegration tests (Westerlund, 2007) are reported in Appendix E. All four statistics, the group-mean G_a and G_t and the panel P_a and P_t, are applied to the main quantity specifications. G_a and P_t reject the null of no cointegration at the 1% level; G_t is insignificant (expected with $\bar{T} \approx 16$ per country, which provides limited power for the individual-country t-statistic); and P_a is borderline. The pattern is consistent with panel-level cointegration in the main quantity KPF, supporting AMG and CCEMG as the preferred basis for inference and, crucially, validating the PMG error-correction framework of Section 4.4, whose consistency is conditional on a cointegrating relationship.

The *quality* specification raises a distinct and more challenging problem. The logit-transformed top-1% citation share is I(0) while the regressors remain I(1), so no cointegrating relationship between the dependent variable and the regressors can be established by construction. This integration-order mismatch means that FE, AMG, CCEMG, and standard PMG, all of which are designed for settings in which the dependent variable and regressors share the same integration order, or in which a cointegrating relationship exists, are not well suited to the quality equation. Whether they produce formally inconsistent estimates in this setting is a theoretical question that goes beyond the scope of this paper; what is clear empirically is that they produce highly unstable results: $\hat{\sigma}_q$ ranges from -10.3 (AMG) to $+1.3$ (FE) across estimators in the preferred Spec Q-B, with standard PMG returning -0.39 ($t = -1.02$). This instability is an empirical signal of estimators applied outside their designed domain.

PMG+CSD (PMG augmented with cross-sectional mean correction) is better motivated for this setting. By absorbing cross-sectional dependence through cross-sectional mean augmenters in the *short-run* equation rather than relying on a long-run cointegrating structure, PMG+CSD does not require a cointegrating relationship between the I(0) dependent variable and the I(1) regressors. This does not constitute a formal theoretical guarantee of consistency in the I(0)/I(1) case (such a result has not, to our knowledge, been established for PMG+CSD in this exact setting) but it removes the specific source of misspecification that afflicts the other estimators, and it produces a stable, precisely estimated result: $\hat{\sigma}_q = 2.188$ ($t = 11.16$), consistent across all reasonable perturbations of the specification. We treat this as the preferred estimator for the quality equation on methodological grounds and on the basis of its empirical stability relative to the alternatives, while acknowledging that the theoretical properties of any estimator under I(0)/I(1) mixed integration in a cross-sectionally dependent panel remain an open research question.

4.3 Estimators

We employ four primary estimators, each addressing a different set of issues:

4.3.1 Two-Way Fixed Effects (FE)

The FE estimator within-demeans all variables by country, removing μ_i, and adds year fixed effects λ_t. It imposes slope homogeneity ($\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all i) and assumes cross-sectional independence conditional on the fixed effects. As discussed above, both assumptions are potentially violated in our application; the FE results are accordingly treated as a benchmark and a basis for comparison rather than as the preferred estimate.

4.3.2 Augmented Mean Group (AMG)

The Augmented Mean Group estimator of Bond and Eberhardt (2009) and Eberhardt and Bond (2009) explicitly models the common dynamic process $\mathbf{f}_t$ via a three-step procedure, each step tying the estimation to the global science frontier that generates cross-sectional dependence.

Step 1. A first-differenced pooled OLS regression is estimated across all countries and periods, including year dummy variables $\{d_t\}$:

$$\Delta y_{it} = \sum_t \mu_t \, \Delta d_t + \sum_k \beta_k \, \Delta x_{k,it} + u_{it}. \tag{12}$$

The estimated coefficients $\{\hat{\mu}_t\}$ on the time dummies are interpreted as increments to an unobserved common dynamic process (CDP). The accumulated CDP is:

$$\hat{C}_t = \sum_{s=t_0}^{t} \hat{\mu}_s, \tag{13}$$

which captures the evolving contribution of the global science frontier to each country's knowledge production.

Step 2. Country-specific OLS regressions are estimated in levels, augmenting the specification with a linear time trend and the estimated CDP:

$$y_{it} = \alpha_i + \boldsymbol{\beta}_i' \mathbf{x}_{it} + \lambda_i \, (t - \bar{t}) + \gamma_i \, \hat{C}_t + \varepsilon_{it}. \tag{14}$$

In specifications using the PWT human capital index (Specs B and D), the linear trend $\lambda_i(t - \bar{t})$ is omitted. The human capital index is itself near-deterministically trending within each country, creating near-perfect collinearity with the explicit trend term that generates explosive country-level coefficients. Dropping the trend is consistent with the recommendation in Eberhardt and Teal (2011) for highly persistent regressors.

Step 3. The AMG estimate is the simple mean of the country-specific slope vectors:

$$\hat{\boldsymbol{\beta}}_{\text{AMG}} = N^{-1} \sum_{i=1}^{N} \hat{\boldsymbol{\beta}}_i, \qquad \widehat{\text{SE}}(\hat{\boldsymbol{\beta}}_{\text{AMG}}) = \frac{s(\hat{\boldsymbol{\beta}}_i)}{\sqrt{N}}, \tag{15}$$

where $s(\hat{\boldsymbol{\beta}}_i)$ is the cross-country standard deviation. Inference is based on a t-distribution with $N - 1$ degrees of freedom.

To ensure robustness of the mean-group average, we exclude country-level estimates lying more than three interquartile ranges from the cross-country median for any coefficient, thereby preventing estimation failures in individual country regressions from distorting the panel average. The trimmed country count is reported alongside each set of estimates.[16]

4.3.3 Common Correlated Effects Mean Group (CCEMG)

The CCEMG estimator of Pesaran (2006) controls for cross-sectional dependence via a different route. Instead of estimating the common factor directly, it proxies $\mathbf{f}_t$ non-parametrically by augmenting each country's regression with the cross-sectional means of all variables:

$$y_{it} = \alpha_i + \boldsymbol{\beta}_i' \mathbf{x}_{it} + \boldsymbol{\delta}_i' \bar{\mathbf{z}}_t + \varepsilon_{it}, \tag{16}$$

where $\bar{\mathbf{z}}_t = (\bar{y}_t, \, \bar{\mathbf{x}}_t')'$ and $\bar{y}_t = N^{-1} \sum_i y_{it}$. Under the assumption that the number of unobserved common factors does not exceed the number of observables included in $\bar{\mathbf{z}}_t$, the

[16] Standard implementations such as Stata's `xtmg` command (Eberhardt, 2012) do not apply this trim by default; replication code implementing it is provided with the paper.

term $\boldsymbol{\delta}_i'\bar{\mathbf{z}}_t$ asymptotically spans the same space as $\boldsymbol{\lambda}_i'\mathbf{f}_t$, rendering the augmented regression consistent. Country-specific estimates $\hat{\boldsymbol{\beta}}_i$ are obtained by OLS applied to equation (16), and the CCEMG estimate is the cross-country mean, computed as in equation (15).

The CCEMG estimator is consistent under cross-sectional dependence, slope heterogeneity, and mixed-order regressors (some I(0), some I(1) on the right-hand side) when the dependent variable is I(1), provided $(N, T) \to \infty$ jointly (Kapetanios et al., 2011; Pesaran, 2006). This result applies to the quantity specification. It does not extend to the quality specification, where the problem is categorically different: the dependent variable itself is I(0) while all regressors are I(1), a setting for which the CCEMG consistency result was not derived and for which PMG+CSD is the preferred estimator (Section 4.2). In our application, $T = 20$ is short relative to the number of regressors augmented by cross-sectional means, so the CCEMG country-level estimates may have finite-sample imprecision; this is mitigated by the same outlier-trimming procedure applied to the AMG. Both AMG and CCEMG are presented alongside FE to triangulate findings across the three estimators.

4.3.4 Pooled Mean Group (PMG)

The Pooled Mean Group estimator of Pesaran et al. (1999) provides a formal test of whether the long-run KPF coefficients are homogeneous across countries. PMG constrains the long-run slopes to be equal across all groups (pooled) while allowing short-run adjustment dynamics and country-specific error-correction speeds π_i to vary freely. It is more efficient than the fully heterogeneous MG when the pooling restriction holds, but long-run estimates are inconsistent if it fails. Validity is assessed via the Hausman test (Table 7) comparing MG and PMG long-run vectors; failure to reject H_0 implies the pooled restriction is appropriate. We implement both standard PMG and PMG+CSD (augmented with cross-sectional means of all variables in the short-run equation, analogously to CCEMG's CSD correction). This combination, pooling the long-run restriction across countries while augmenting the short-run equations with cross-sectional means to address cross-sectional dependence, is related to the CS-ARDL framework of Chudik et al. (2016) and to the cross-sectionally augmented error-correction model (CS-ECM) discussed in Ditzen (2021), both of which recover long-run coefficients in heterogeneous panels under cross-sectional dependence. The distinguishing feature of our implementation is the imposition of pooled (homogeneous) long-run slopes via the PMG iterative WLS step, rather than estimating country-specific long-run coefficients as in the mean-group CS-ARDL. Results appear in Table 7 and are discussed in Section 5.2.

The world knowledge stock and international spillovers: implications for the estimators. Because K_{t-1}^{world} is invariant across countries within each year, it is exactly collinear with the year fixed effects in FE, the Common Dynamic Process in AMG, and

the cross-sectional mean augmenters in CCEMG: absorbed by whichever global-trend control each estimator employs, and therefore unidentifiable as a regressor. The rest-of-world stock $K_{t-1}^{-i} = K_{t-1}^{world} - K_{i,t-1}$ avoids exact collinearity through the log nonlinearity (equation (6)), but the within-year cross-sectional standard deviation of $\ln K_{t-1}^{-i}$ is only 0.036, concentrated among the five or six largest science economies. The implications for what each estimator can identify can be summarised directly. **FE** is the most informative estimator for the RoW specification: double-demeaning removes almost all variation in $\ln K_{t-1}^{-i}$ (which tracks the global trend), leaving a small orthogonal cross-sectional residual that identifies $\hat{\phi}_2$ separately from $\hat{\phi}_1$. **AMG** and **CCEMG** by contrast generate near-identification failures because AMG's Common Dynamic Process and CCEMG's cross-sectional mean augmenters already absorb the global knowledge trend as a common factor. Adding $\ln K_{t-1}^{-i}$, which for most countries is nearly identical to $\ln K_{t-1}^{world}$, creates near-collinearity with the internal global-trend controls, and the coefficient vector shifts to absorb this near-multicollinearity. The identification issues for each estimator, and why the resulting estimation results we observe across AMG and CCEMG we observe is confirmatory, are derived in full in Appendix D, alongside the empirical results (Table 10). Appendix D.1 derives testable predictions from the identification: **FE**: $\phi_1 > 0$, $\phi_2 > 0$, separable; **AMG**: $\phi_1 \to 0$, $\phi_2 > 0$ and inflated; **CCEMG**: $\phi_1 \to 0$, $\phi_2 > 0$ and explosive.

4.4 Robustness Checks

We assess the robustness of the main findings along six dimensions:

1. *Knowledge stock depreciation rate.* The baseline $\rho_K = 0.10$ is re-estimated using $\rho_K = 0.05$ and $\rho_K = 0.15$. Variation in ρ_K alters the level and persistence of the knowledge stock, providing a check on whether the ϕ estimate is sensitive to the assumed speed of knowledge obsolescence.

2. *Product variety proxy.* The patent stock is replaced successively by (a) the annual flow of patent applications; (b) real GDP (PPP-adjusted, from PWT 11.0); and (c) total population. These alternatives address the concern that the patent stock is not a credible proxy for the Jones (1995) product variety concept Q and may instead reflect innovative capacity.

3. *Knowledge quality.* The quality specification replaces the quantity-based article count with the logit-transformed top-1% citation share. If the KPF captures not merely the production of articles but the production of high-impact scientific output, the same input variables should predict citation quality conditional on the knowledge stock.

4. *Rest-of-world knowledge spillover.* The baseline specification is augmented with

$\ln K_{t-1}^{-i}$ alongside $\ln K_{i,t-1}$, testing whether results conform to the identification predictions of Appendix D.

5. *Long-run homogeneity (PMG).* The Pooled Mean Group estimator formally tests whether long-run KPF coefficients are homogeneous across countries via a Hausman comparison of MG and PMG estimates (Table 7).

6. *R&D expenditure (GERD).* A concern is that $\hat{\phi}$ proxies for sustained national investment in R&D rather than a structural compounding mechanism. We augment the baseline quantity and quality specifications with $\ln(\text{GERD}_{it})$ (log of absolute real R&D expenditure, constructed as $\ln(g_{it} \times Y_{it})$ where g_{it} is the WDI GERD-to-GDP ratio (World Bank, 2026) and Y_{it} is PWT real GDP) to test whether the knowledge stock elasticity is displaced (Table 11, discussed in Sections 5.2 and 5.3).

5 Results

5.1 Outline of Estimation Results Presentation

Tables 4 through 10 report results for the quantity of scientific output specifications. Tables 4 and 5 report the baseline KPF estimations for both the above-median and non-zero samples. Robustness tests proceed as follows: Table 6 examines sensitivity to depreciation rate assumptions ($\rho_K \in \{0.05, 0.10, 0.15\}$); Table 7 presents the PMG long-run homogeneity check; Table 8 explores sensitivity to product variety specification; Table 10 investigates the impact of how the rest of the world is treated in estimation (on which see also Appendix D). Results are discussed in detail in Section 5.2.

The quality of scientific output specification results are presented in Table 12. Results are discussed in Section 5.3.

While our results favor specifications with KPF input elasticities that are homogeneous across countries, Section 5.2 presents some brief reflections on country-specific evidence to emerge from the AMG and CCEMG estimators for the quantity of scientific output evidence (the quality of scientific evidence rests on PMG+CSD estimation in which input elasticity homogeneity is strict).

5.2 Discussion: Quantity of Scientific Output Results

The discussion below walks through each estimated parameter in turn: baseline results, contrasted with robustness considerations, with the AMG/CCEMG prioritisation principle and the PMG results addressed where they bear directly on each parameter.

$\hat{\phi}$: knowledge stock elasticity — baseline. The most robust finding across all quantity specifications, estimators, and samples is a positive and highly significant knowledge

Table 4: KPF Estimates: Above-Median and Non-Zero Samples, Whole-Count Articles

	Above-median Spec A H = Researchers ($N = 81$)			*Above-median* Spec B H = PWT hc ($N = 81$)			*Non-zero* Spec B† H = PWT hc ($N = 137$)	
	FE	AMG	CCEMG	FE	AMG	CCEMG	AMG	CCEMG
$\hat{\sigma}$: $\ln(H_A)$	0.385***	0.042	0.126*	0.166	2.300***	1.056	4.165***	3.638***
	(0.033)	(0.038)	(0.067)	(0.656)	(0.802)	(0.910)	(0.871)	(1.042)
$\hat{\gamma}$: $\ln(Q)$	0.116***	0.218	0.011	0.046	−0.048	−0.011	−0.037	0.024
	(0.020)	(0.150)	(0.118)	(0.045)	(0.103)	(0.098)	(0.078)	(0.080)
$\hat{\phi}$: $\ln(K_{t-1})$	0.044***	0.419*	0.549***	0.646***	0.762***	0.751***	0.561***	0.514***
	(0.007)	(0.212)	(0.094)	(0.014)	(0.116)	(0.107)	(0.089)	(0.095)
Countries in MG	81	57	52	81	72	68	111	106
Observations		1,032			1,383		2,033	

Notes: Dependent variable: ln(annual whole-count articles). FE = two-way fixed effects. AMG = Augmented Mean Group (Bond & Eberhardt, 2009; Eberhardt & Bond, 2009). CCEMG = Common Correlated Effects Mean Group (Pesaran, 2006). Standard errors in parentheses. ***, **, *: 1%, 5%, 10%. AMG and CCEMG trim country-level estimates $> 3\times$ IQR before mean-grouping; trimmed N in 'Countries in MG'. K_{t-1} = lagged knowledge stock (PIM, $\rho_K = 0.10$).

†*Non-zero sample* covers 181 countries with positive article output; 137 have complete data for Spec B over the full period. FE non-zero results are omitted: the FE estimate of $\hat{\sigma}$ is near zero (0.033, SE = 0.419, insignificant), as expected given that the PWT hc index has negligible within-country variation over $T = 20$. Specs A, C, D for the non-zero sample are qualitatively consistent and available in Online Appendix B.3.

Table 5: KPF Estimates: Above-Median and Non-Zero Samples, Fractional-Count Articles

	Above-median Spec C H = Researchers ($N = 81$)			*Above-median* Spec D H = PWT hc ($N = 81$)			*Non-zero* Spec D† H = PWT hc ($N = 120$)	
	FE	AMG	CCEMG	FE	AMG	CCEMG	AMG	CCEMG
$\hat{\sigma}$: $\ln(H_A)$	0.388***	0.038	0.069	0.075	2.491***	1.352	3.790***	3.321***
	(0.036)	(0.038)	(0.057)	(0.659)	(0.873)	(1.001)	(0.722)	(0.954)
$\hat{\gamma}$: $\ln(Q)$	0.110***	0.057	0.203	0.050	0.022	0.162	−0.024	0.095
	(0.021)	(0.152)	(0.170)	(0.045)	(0.118)	(0.116)	(0.099)	(0.092)
$\hat{\phi}$: $\ln(K_{t-1})$	0.045***	0.733***	0.500***	0.647***	0.652***	0.505***	0.462***	0.293**
	(0.008)	(0.182)	(0.105)	(0.014)	(0.127)	(0.122)	(0.115)	(0.119)
Countries in MG	81	55	50	81	71	67	120	123
Observations		1,032			1,383		2,033	

Notes: Dependent variable: ln(annual fractional-count articles). See notes to Table 4.
†*Non-zero Spec D*: 181 countries; 120 (AMG) and 123 (CCEMG) have complete data after trimming. FE non-zero result omitted (PWT hc index has negligible within-country variation over $T = 20$). Non-zero Spec C available in Online Appendix B.3.

Table 6: Robustness – Alternative Knowledge Stock Depreciation Rates (Spec B, Above-Median)

	$\rho_K = 0.05$		$\rho_K = 0.10$ (baseline)		$\rho_K = 0.15$	
Estimator	$\hat{\sigma}$	$\hat{\phi}$	$\hat{\sigma}$	$\hat{\phi}$	$\hat{\sigma}$	$\hat{\phi}$
FE	0.0994	0.6399***	0.1664	0.6457***	0.2254	0.6500***
	(0.6610)	(0.0142)	(0.6559)	(0.0142)	(0.6526)	(0.0141)
AMG	2.6629***	0.6683***	2.2996***	0.7620***	3.0878***	0.7716***
	(0.9758)	(0.1331)	(0.8017)	(0.1164)	(0.8623)	(0.1055)
CCEMG	0.3356	0.7460***	1.0556	0.7508***	1.5282*	0.7371***
	(1.0555)	(0.1177)	(0.9102)	(0.1073)	(0.8410)	(0.0890)

Notes: Dependent variable ln(whole-count articles); H_A = PWT hc index; N = 81 countries. Patent stock included in all specifications (coefficients omitted for brevity). Standard errors in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table 7: Pooled Mean Group (PMG) Robustness Check — Long-Run Coefficients and Hausman Test (Spec B, Above-Median)

	FE	AMG	CCEMG	PMG	PMG+CSD
$\hat{\sigma}$ (Human capital)	0.214	2.927***	1.150	3.279***	1.825***
	(0.618)	(0.791)	(0.842)	(1.230)	(0.700)
$\hat{\gamma}$ (Patent stock)	0.046	0.076	−0.007	0.041	0.061
	(0.044)	(0.087)	(0.096)	(0.116)	(0.066)
$\hat{\phi}$ (Knowledge stock)	0.646***	0.715***	0.667***	0.503***	0.652***
	(0.014)	(0.119)	(0.115)	(0.034)	(0.020)
Countries (N)	81	73	74	79	79
Hausman test: H_0: long-run homogeneity valid ⇒ PMG preferred					
AMG vs PMG	$H = 3.16$, $df = 3$, $p = 0.367$ ⇒ fail to reject: PMG valid				
CCEMG vs PMG+CSD	$H = 3.07$, $df = 3$, $p = 0.381$ ⇒ fail to reject: PMG+CSD valid				
Error-correction speeds (PMG, country-specific π_i)					
% negative (stabilising)	77%				
Median (negative only)	−0.457 (implied half-life: 1.5 years)				
Inter-quartile range	[−0.660, −0.038]				

Notes: Dependent variable: ln(whole-count articles); H_A = PWT hc index; Q = patent stock; K_{t-1} = lagged knowledge stock ($\rho_K = 0.10$). FE = two-way fixed effects. AMG = Augmented Mean Group. CCEMG = Common Correlated Effects MG. PMG = Pooled Mean Group (Pesaran et al., 1999); long-run θ estimated by iterative WLS, short-run dynamics and error-correction speeds country-specific. PMG+CSD = PMG augmented with cross-sectional means of all variables in the short-run equation (CSD correction analogous to CCEMG). The Hausman test compares the MG estimator (AMG or CCEMG) with the corresponding pooled estimator (PMG or PMG+CSD); under H_0, the long-run homogeneity restriction is valid and PMG is more efficient. Both tests are passed at conventional significance levels. Standard errors in parentheses. ***, **, *: significance at 1%, 5%, 10%.

Table 8: Robustness — Alternative Product Variety (Q) Proxies (Spec B, Above-Median)

	FE			AMG			CCEMG		
Q proxy	$\hat{\sigma}$	$\hat{\gamma}$	$\hat{\phi}$	$\hat{\sigma}$	$\hat{\gamma}$	$\hat{\phi}$	$\hat{\sigma}$	$\hat{\gamma}$	$\hat{\phi}$
Patent stock (baseline)	0.1664	0.0462	0.6457***	2.2996***	−0.0475	0.7620***	1.1502	−0.0072	0.6672***
	(0.6559)	(0.0447)	(0.0142)	(0.8017)	(0.1029)	(0.1164)	(0.8420)	(0.0959)	(0.1154)
Patent applications (flow)	−0.0726	0.1471***	0.6464***	2.3315***	0.0388	0.6904***	1.1020	0.0626**	0.7856***
	(0.6506)	(0.0538)	(0.0141)	(0.8577)	(0.0264)	(0.0776)	(0.8124)	(0.0252)	(0.0733)
Real GDP	0.4881	0.2666***	0.6470***	2.8305***	−0.0054	0.6226***	−0.1101	−0.0594	0.4465***
	(0.5583)	(0.0591)	(0.0132)	(0.7569)	(0.0608)	(0.0705)	(1.3833)	(0.0772)	(0.1195)
Population	−0.3623	1.0230***	0.6507***	1.7803*	−0.6101	0.6931***	−0.3743	2.5909*	0.5661***
	(0.5857)	(0.2241)	(0.0132)	(0.9572)	(1.0493)	(0.1105)	(1.6452)	(1.3000)	(0.2039)
Countries	81–83			72–77			72–74		

Notes: Dependent variable: ln(whole-count articles); H_A = PWT hc index; K_{t-1} = lagged knowledge stock ($\rho_K = 0.10$). Each row substitutes the column header variable for $\hat{\gamma}$: ln(Q) as the product-variety proxy Q. Columns show $\hat{\sigma}$ (HC elasticity), $\hat{\gamma}$ (Q elasticity), $\hat{\phi}$ (knowledge stock elasticity). Country count varies by Q proxy coverage. Standard errors in parentheses. The central finding — $\hat{\phi} > 0$ significant across all proxies and all estimators — is robust throughout. The Q elasticity $\hat{\gamma}$ is small and sign-unstable, consistent with the patent stock not being a reliable empirical measure of the Jones (1995) product-variety concept. ***, **, *: significance at 1%, 5%, 10%.

stock elasticity $\hat{\phi}$ (Tables 4 and 5). In Spec B (PWT human capital index, whole-count output, above-median sample), the preferred quantity specification, combining the broadest country coverage with the quality-adjusted human capital measure that most closely matches theory, $\hat{\phi}$ is estimated at 0.646 (FE), 0.762 (AMG), and 0.751 (CCEMG). The fractional-count specifications (Specs C and D, Table 5) deliver closely consistent estimates: Spec D (PWT hc) yields $\hat{\phi} = 0.647$ (FE), 0.652 (AMG), and 0.505 (CCEMG), while Spec C (researcher count) yields $\hat{\phi} = 0.045$ (FE), 0.733 (AMG), and 0.500 (CCEMG); the same attenuation in FE under the researcher-count specification noted below. Moreover, the knowledge stock elasticity $\hat{\phi}$ is remarkably stable across the two samples (above-median: 0.646–0.751; non-zero: 0.514–0.644), confirming that compounding is a structural feature of knowledge production at all stages of development, not an artefact of the above-median sample selection.[17]

These values lie within the semi-endogenous range ($0 < \phi < 1$) but substantially above prior estimates from patent-based studies ($\hat{\phi} \approx 0.3$–0.5; Jones, 1995a), indicating that for academic science output, the compounding of accumulated knowledge is a more powerful mechanism than for commercial innovation. This finding is consistent with the non-rival, public-good character of peer-reviewed science.

At the *country* level, a small number of mean-group estimates exceed the Romer threshold ($\hat{\phi}_i > 1$), constituting locally strong Schumpeterian results, but these are exceptions to the dominant homogeneity finding rather than evidence against it. The panel average ($\bar{\hat{\phi}} \approx 0.76$) reflects the semi-endogenous regime, with the exceptions discussed in the efficiency rankings paragraph below.

$\hat{\phi}$: knowledge stock elasticity — robustness. The $\hat{\phi}$ estimates are fully robust to the depreciation rate (Table 6): across $\rho_K \in \{0.05, 0.10, 0.15\}$, the FE estimate ranges from 0.640 to 0.650, the AMG estimate from 0.668 to 0.772, and the CCEMG estimate from 0.737 to 0.746. This stability, spanning a three-fold range in the assumed depreciation rate, is reassuring given the absence of strong prior evidence on the appropriate rate of knowledge obsolescence in the bibliometric context, and confirms that the finding is not an artefact of the perpetual inventory assumptions.[18]

[17] We focus the discussion below on the above-median sample as the more informative specification for established science economies; the non-zero sample results are interpreted alongside the above-median findings where appropriate.

[18] Note that we also tested for the absorptive capacity hypothesis by including a multiplicative HC×K interaction term. The test yields a negative and multicollinear result: the interaction coefficient is significant (FE: -0.817^{***}; AMG: -2.424^{**}) but negative, and adding it inflates the main effects dramatically ($\hat{\sigma}$ rising from 0.21 to 7.40 in FE; $\hat{\phi}$ from 0.65 to 1.54): a signature of near-perfect multicollinearity among $\ln H$, $\ln K$, and their product within a $T = 20$ time series. The test cannot distinguish the multiplicative from the additive structure given the available time dimension; the additive specification is the appropriate basis for inference. Whether HC and knowledge stocks are truly multiplicative complements in the Cohen–Levinthal sense remains an open empirical question better addressed with longer time-series data at the field level.

In the non-zero sample (Table 4 (non-zero panel)), $\hat{\phi}$ is 0.644 (FE), 0.561 (AMG), and 0.514 (CCEMG), below the above-median sample but of the same order, confirming that knowledge-stock compounding is operative at all stages of scientific development, not only among established science powers.

A note on the FE–AMG/CCEMG divergence in the researcher-count specifications (Specs A and C): the FE estimate of $\hat{\phi}$ drops to approximately 0.04–0.05 in these specifications, substantially below the AMG/CCEMG estimates of 0.42–0.73. This pattern arises because FE's common-slope, within-country identification absorbs a portion of the knowledge stock effect into the year fixed effects when the researcher headcount and the knowledge stock are both trending at similar rates within countries. AMG and CCEMG, which allow heterogeneous loadings on the global frontier process, recover the fuller cross-country variation in the compounding advantage. The Spec B FE estimate (0.646) does not exhibit this attenuation because the PWT human capital index trends very differently from the knowledge stock within countries, breaking the collinearity that suppresses $\hat{\phi}$ in the researcher-count specifications.

The Pooled Mean Group estimator (Table 7, quantity specifications) provides the most precise confirmation of the baseline findings. PMG+CSD returns $\hat{\phi} = 0.652^{***}$ (SE = 0.020), within rounding of both the FE (0.646) and CCEMG (0.667) estimates, but with standard errors approximately six times smaller than CCEMG (0.115). The Hausman test fails to reject long-run homogeneity (CCEMG vs PMG+CSD: $H = 3.07$, $p = 0.381$), confirming that the pooling restriction is valid and that the heterogeneous-panel mean-group estimates are averaging a genuine structural common relationship rather than a noisy mixture of country-specific patterns.

The take-away: $\hat{\phi} \approx 0.65$–0.76 is a robust, stable feature of cross-country academic knowledge production dynamics, confirmed across three estimator families (FE, MG, PMG), three depreciation rates, two output measures, two human capital proxies, and two country samples. The semi-endogenous interpretation, knowledge compounding is real, substantial, and closer to the Romer threshold than prior technological innovation estimates suggest, is a central empirical finding of the paper.

$\hat{\sigma}$: human capital elasticity — baseline. The estimated human capital elasticity $\hat{\sigma}$ varies substantively across specifications and estimators, reflecting the different ways in which each handles the identification challenge posed by the human capital variable (Tables 4 and 5).

In the *researcher-count* specifications (Specs A and C), FE returns $\hat{\sigma} \approx 0.38$–$0.39$, strongly significant at conventional levels and consistent with the cross-country estimates of Guellec and van Pottelsberghe de la Potterie (2004). The AMG and CCEMG estimates are smaller (0.04–0.13) and at best marginally significant. The result that FE, the estimator most prone to attenuation, records a larger $\hat{\sigma}$ than the heterogeneous-panel

estimators, is not paradoxical once the within-vs-between variation in researcher stocks is understood: total researcher counts grow relatively slowly and uniformly within countries over the 20-year window, generating sufficient within-country variation for FE to identify a positive $\hat{\sigma}$. AMG and CCEMG, however, control for the common global frontier process across all countries simultaneously; since researcher-count trends are strongly correlated with the global science frontier (countries that grow their researcher stocks fastest are precisely those converging most rapidly toward the frontier), their internal frontier controls absorb a large fraction of the between-country researcher-count variation, leaving less to identify $\hat{\sigma}$ from. FE exploits residual *within*-country researcher variation that its common-year-effect control does not absorb; AMG and CCEMG absorb more of the cross-country dimension through the global frontier process, and $\hat{\sigma}$ is the casualty.

In the *PWT human capital index* (PWT hc) specifications (Specs B and D), FE estimates of $\hat{\sigma}$ are near zero and entirely insignificant. This is expected: the PWT human capital index is an extremely smooth, near-deterministic function of time within each country (it is built from schooling attainment and returns-to-education estimates, which change very gradually), so country fixed effects remove virtually all identifying within-country variation. AMG, which controls for the common global frontier process but preserves cross-country heterogeneity in the frontier loading, reveals a positive and highly significant human capital elasticity: $\hat{\sigma} \approx 2.30^{***}$ (Spec B) and 2.49^{***} (Spec D). These substantially exceed unity, suggesting super-proportional returns to human capital quality in knowledge production. CCEMG, which imposes fewer restrictions on the factor structure, returns $\hat{\sigma} \approx 1.06$ (Spec B) and 1.35 (Spec D), insignificant but directionally consistent with the AMG estimates.

The fractional-count specifications (Table 5) corroborate the PWT hc pattern. Spec D (fractional count, PWT hc) returns $\hat{\sigma} = 0.075$ (FE, insignificant), 2.491^{***} (AMG), and 1.352 (CCEMG), almost identical to Spec B, confirming the finding is not sensitive to the article-count weighting convention. Spec C (fractional count, researcher count) returns $\hat{\sigma} = 0.388^{***}$ (FE), 0.038 (AMG), and 0.069 (CCEMG), again matching the researcher-count pattern in Spec A.

In the non-zero sample (Table 4, non-zero panel), AMG returns $\hat{\sigma} = 4.17^{***}$ and CCEMG $\hat{\sigma} = 3.64^{***}$, substantially larger than the above-median estimates. This is suggestive of extensive-margin dynamics of science entry: at the transition from zero to positive science output, human capital quality is the critical binding constraint, and its marginal returns are correspondingly large.

$\hat{\sigma}$: human capital elasticity — robustness and take-away. Across the depreciation robustness checks (Table 6), $\hat{\sigma}$ is less stable than $\hat{\phi}$, reflecting the genuine imprecision of the HC elasticity in the AMG specification. The AMG estimate ranges from 2.26^{***} to 3.09^{***} across $\rho_K \in \{0.05, 0.10, 0.15\}$; always positive and significant, but with

a coefficient of variation substantially larger than $\hat{\phi}$'s. The CCEMG estimate remains directionally consistent but crosses the significance boundary at some depreciation rates. This instability does not undermine the qualitative finding; it reflects the inherent difficulty of separately identifying the level effect of a slowly-varying quality index from the compounding effect of the accumulated knowledge stock in a $T = 20$ panel.

PMG+CSD (Table 7) returns $\hat{\sigma} = 1.825^{***}$ (SE = 0.700) for the quantity specification, falling between AMG and CCEMG and providing the most efficient estimate under the pooling restriction. The Hausman test confirms the pooling restriction is valid, so PMG+CSD is the preferred basis for inference: human capital quality significantly raises knowledge production, with an elasticity of approximately 1.8.

The take-away: human capital quality ($\hat{\sigma} \approx 1.8$–2.3, preferred AMG and PMG+CSD estimates) is a strong and significant driver of academic knowledge production. Its elasticity exceeds that of the knowledge stock in the above-median sample, suggesting that for established science economies, the quality of the research workforce is at least as important as the depth of the inherited knowledge base in determining current output. The very large non-zero-sample estimates ($\hat{\sigma} \approx 3.6$–4.2) point to human capital quality as the binding constraint for countries on the extensive margin of science participation.

$\hat{\gamma}$: patent stock elasticity — product variety proxy. The estimated elasticity with respect to the patent stock $\hat{\gamma}$ is generally small, sign-unstable, and statistically insignificant in AMG and CCEMG across all quantity specifications (Tables 4 and 5). This instability is confirmed in the product-variety robustness check (Table 8): substituting patent applications (flow), real GDP, or population for the patent stock changes the point estimate substantially while leaving $\hat{\phi}$ virtually unchanged, confirming that the instability lies with the patent stock as a proxy rather than with the product-variety concept itself.[19] The GDP and population alternatives yield positive and more stable FE estimates (GDP: $\hat{\gamma} = 0.27^{***}$; population: $\hat{\gamma} = 1.02^{***}$), suggesting that broader proxies for market scale capture the product-variety concept better than commercial patents in a science production context.

The deeper issue is conceptual: the Jones (1995) product-variety construct was designed for commercial innovation, not academic science. A more natural academic proxy, for instance the number of active Web of Science subject categories per country-year, is to be best of our knowledge not currently available in panel form. The consistent finding is that $\hat{\phi}$ is entirely robust to the choice of Q proxy (Table 8), while $\hat{\gamma}$ is not: the knowledge-stock compounding finding does not depend on measuring product variety correctly, while

[19]The patent stock may simply not be a reliable empirical proxy for the product variety construct Q in the science production context. The instability pattern we report is consistent with the interpretation that the patent stock in cross-country data captures aspects of country size and industrial structure rather than the theoretical concept of the number of scientific lines of enquiry over which new knowledge must be distributed. The failure to identify a credible Q estimate is a genuine limitation, consistent with the more general challenge of measuring product variety empirically noted by Fedderke and Dong (2026).

the product variety finding is itself inconclusive given available data.

Implications for the extensive and intensive margins of knowledge production. The comparison of above-median and non-zero samples yields an interpretively important contrast. In the non-zero sample (Table 4 (non-zero panel)), the human capital elasticity in AMG and CCEMG is substantially larger (4.2 and 3.6) than in the above-median sample (2.3 and 1.1). The implication is that science system maturity is associated with a shift in the relative engine of output growth, from human capital quality toward knowledge stock compounding, though the absolute return to human capital remains large throughout the distribution. This suggests that at the extensive margin (the entry of peripheral countries into science) human capital quality is the critical binding constraint, and its returns are correspondingly large. At the intensive margin (among already science-active countries, eg. our above-median sample) the human capital elasticity remains substantial (AMG: $\hat{\sigma} \approx 2.3$) and continues to be an important driver of output alongside the knowledge stock. What changes across the distribution is not the importance of human capital *per se* but its relative weight compared to knowledge stock compounding: the compounding channel ($\hat{\phi}$) gains relative importance as the knowledge base deepens, while human capital quality remains a significant independent contributor at all stages. This decomposition is consistent with, and suggestive of, a two-stage model of science development: human capital quality drives the initial establishment of a research base, after which cumulative knowledge advantage becomes the primary driver of output growth.

Country fixed effects as national innovation system quality. The country fixed effects $\hat{\mu}_i$ from the FE specification encode the residual article output not explained by measured inputs, what Nelson (1993) terms the *national innovation system* (NIS) quality: the depth of research infrastructure, the effectiveness of science governance, the strength of university-industry linkages, and the open-science norms that raise the productivity of any given combination of human capital, patent activity, and accumulated knowledge. Despite the limitations of the FE estimator that we have detailed, given the intuitive appeal of the $\hat{\mu}_i$ interpretation, especially given the discussion of equation (1) and Section 2, we consider the implications of our results on $\hat{\mu}_i$. Table 9 reports the estimated $\hat{\mu}_i$ for the fifteen focus economies, ranked from highest to lowest.

Three patterns stand out, read together with the intercept column in Table 9 rather than $\hat{\mu}_i$ alone. First, and surprisingly, China records the largest $\hat{\mu}_i$ in the sample (+1.26, intercept 2.78), more than 0.38 log-points above the United States (+0.88, intercept 2.40): conditional on its human capital index, patent stock, and accumulated knowledge, China produces substantially more articles than the estimated structural relationship predicts. This aligns with Fu (2015)'s account of deliberate state mobilisation of science capacity through targeted funding programmes and institutional incentives. The result implies that

Table 9: Country Fixed Effects ($\hat{\mu}_i$) as NIS Quality, FE Spec B (Above-Median)

ISO	Country	$\hat{\mu}_i$	Intercept†	Interpretation
CHN	China	+1.257	2.779	State-directed research expansion; 985/211 funding programmes
USA	United States	+0.877	2.399	Elite research university system; NIH/NSF infrastructure
IND	India	+0.842	2.364	IIT/IISc system; large diaspora research network
BRA	Brazil	+0.593	2.115	FAPESP/CNPq funding base; USP cluster
GBR	United Kingdom	+0.577	2.099	Russell Group depth; high international collaboration
DEU	Germany	+0.513	2.035	Max Planck/Fraunhofer dual-track system
ITA	Italy	+0.483	2.005	Strong basic science tradition; CNR network
CAN	Canada	+0.467	1.989	NSERC/CIHR system; US proximity spillover
FRA	France	+0.445	1.968	CNRS centralised system; Grandes Écoles research
TUR	Türkiye	+0.450	1.972	TUBITAK; rapid institutional expansion
KOR	South Korea	+0.397	1.919	POSTECH/KAIST system; chaebôl R&D linkages
JPN	Japan	+0.267	1.789	Mature but relatively closed NIS; limited fresh talent inflow
MEX	Mexico	+0.179	1.701	UNAM/CONACYT base; constrained by brain drain
ZAF	South Africa	+0.149	1.671	NRF system; thin base relative to inputs
ISR	Israel	−0.169	1.354	Concentrated excellence over breadth; thin mass
Panel reference				
–	Panel mean (81 countries)	−0.016	1.522	Grand-mean intercept = 1.522
–	Panel SD	+0.434	–	

Notes: $\hat{\mu}_i$ estimated from two-way FE Spec B (PWT hc index, above-median sample, $N = 81$), normalised to sum to zero across the mean-group sample. Fixed effects are identified as within-country mean residuals after partialling out two-way fixed effects. †Intercept = grand-mean intercept (1.522) $+\hat{\mu}_i$: the country's full implied constant term, i.e. predicted ln(articles) when all regressors are at their panel-mean level. Values of $\hat{\mu}_i$ reflect conditional output not explained by measured H, Q, K: positive = higher article output than measured inputs predict; negative = lower. **Israel's negative $\hat{\mu}_i$ does not indicate weak output**: its intercept (1.354) remains well above zero and exceeds 13 of the 81 countries in the mean-group sample; the negative sign only means Israel's article output, conditional on its inputs, is modestly below the *panel average* relationship, consistent with a research system built on a small number of internationally prominent fields rather than broad-based volume. Interpretations are indicative, not causal.

China's advantage in the efficiency rankings (Section 5.2) has two distinct components: a structural compounding advantage ($\hat{\phi}_i \approx 1.09$) and a baseline NIS advantage ($\hat{\mu}_i = +1.26$).

Second, the United States ranks second in NIS quality ($\hat{\mu}_i = +0.88$, intercept 2.40), reflecting the depth of the post-WWII research university and federal agency system that is not fully captured by the PWT human capital index. The US's low composite efficiency rank (sixth among G8 in the lollipop charts of Online Appendix B; see also Section 5.2) therefore reflects not NIS weakness but rather human capital saturation: the NIS is high-quality, but the marginal researcher adds little to aggregate output in an already-large workforce.

Third, Israel is the only focus economy with a negative fixed effect ($\hat{\mu}_i = -0.17$). As the table notes make explicit, a negative $\hat{\mu}_i$ is a statement about Israel's position *relative to the panel average*, not a statement about absolute output: Israel's full intercept (1.354) remains positive and exceeds 13 of the 81 countries in the mean-group sample. Israel produces fewer articles than its inputs predict *under the panel-average structural relationship*, not fewer articles in absolute terms. Israel's efficiency advantage comes entirely from the compounding channel ($\hat{\phi}_i \approx 2.25$, robustly positive across all available combinations): it extracts exceptional output from each unit of its *existing* stock, but the institutional architecture generates modestly below-panel-average baseline output for a country of its input endowment, consistent with a science system characterised by extreme concentration in a few world-leading disciplines rather than broad-based research capacity.

Country-specific efficiency rankings. While the PMG results support homogeneity of long-run KPF elasticities across countries, the heterogeneous-panel estimators permit recovery of country-specific coefficients that carry implications for the efficiency of national science systems beyond the fixed-effect NIS quality already documented. The composite efficiency score for country i is $C_i = \hat{\sigma}_i \cdot \bar{H} + \hat{\gamma}_i \cdot \bar{Q} + \hat{\phi}_i \cdot \bar{A}$, where $\bar{H}$, $\bar{Q}$, $\bar{A}$ are the cross-country panel means of the PWT human capital index, patent stock, and knowledge stock respectively. Evaluating at the panel-average input endowment strips out cross-country differences in factor scale, so that differences in C_i reflect differences in elasticities alone — a ranking of production efficiency independent of input size. Ranks are computed within each of the eight specification × estimator combinations (Specs A–D × AMG, CCEMG) and averaged across combinations. Three caveats are essential at the outset. Country-specific elasticities are estimated from $T = 20$ observations per country, so ranges across combinations are wide and pairwise differences between adjacent-ranked countries are rarely distinguishable. Composite scores are sensitive to the panel-mean weights; countries with input endowments far from the mean would rank differently under country-own weighting. And the rankings reflect period-average efficiency over 2003–2022, masking

any time variation.[20]

With those caveats stated, four findings survive the full range of specification-estimator combinations and can therefore be reported with reasonable confidence. First, China's $\hat{\phi}_i \approx 1.09$ exceeds unity in all eight combinations, the only G8 economy that consistently crosses the Romer proportionality threshold, identifying China as the most efficient converter of accumulated knowledge among the major science economies. Second, Japan's composite score is last in the G8 in every combination, consistent with a technologically mature system facing diminishing rather than compounding returns to its knowledge base over this period. Third, among the emerging markets, Türkiye's $\hat{\phi}_i \approx 2.90$ is robustly above unity across all eight combinations (the highest in the full sample of fifteen economies) reflecting exceptional compounding returns on a rapidly growing but relatively young knowledge base. Fourth, South Korea is the only emerging market for which both $\hat{\sigma}_i$ and $\hat{\gamma}_i$ are positive across all combinations, pointing to consistently positive marginal returns to both human capital and patent activity. Beyond these four cases, the min-to-max ranges across combinations are sufficiently wide, and frequently straddle zero, that the individual rankings of the remaining economies should be treated as illustrative of directional tendencies rather than precise efficiency measurements. This conclusion is itself consistent with the parameter homogeneity finding from the PMG analysis: the cross-country distribution of KPF elasticities is tight enough that pooling is valid, and only the most distinctive national experiences generate country-specific estimates that are robustly distinguishable from the panel mean at $T = 20$.

Methodological Reflections on Estimators. The close agreement between AMG and CCEMG for $\hat{\phi}$ throughout, and their systematic divergence for $\hat{\sigma}$ in the PWT hc specifications, is itself informative. Both estimators are consistent under cross-sectional dependence and panel unit roots in the quantity specification. Their key difference is that AMG extracts the common dynamic process via a first-differenced auxiliary regression, while CCEMG augments each country's regression with cross-sectional means, which is more robust under multiple common factors. Both are designed for I(1) systems; neither is the preferred estimator for the quality specification's I(0)/I(1) integration-order mismatch, which is addressed by PMG+CSD (Section 4.2). When they agree closely, as for $\hat{\phi}$, this convergence is evidence that the finding is not an artefact of either estimator's specific identification strategy. When they diverge, as for $\hat{\sigma}$ in Specs B and D, the pattern is specific and informative: AMG returns a large and significant $\hat{\sigma}$ ($\approx$ 2.3–2.5), while CCEMG returns a smaller, insignificant estimate. Rather than defaulting to the more conservative (CCEMG) estimate in this case, we note that the PMG+CSD result

[20]Lollipop charts of the individual elasticities and composite scores for the G8 and emerging market groups, together with full country-specific scatter plots across all eight combinations, are provided in Online Appendix B. The uncertainty bars in those figures show the min-to-max range across the eight combinations, not a conventional standard error, and the range is wide for most countries.

($\hat{\sigma} = 1.825^{***}$) falls between the two and provides an independent, pooling-efficient confirmation that the AMG estimate is not spurious. We therefore treat AMG (corroborated by PMG+CSD) as the basis for inference on $\hat{\sigma}$, while treating AMG–CCEMG agreement as strengthening the evidence for $\hat{\phi} \approx 0.65$–0.76.

The observation that country-specific coefficient estimates do not diverge much across economies motivates a formal efficiency test via the Pooled Mean Group estimator of Pesaran et al. (1999). PMG constrains the long-run coefficients to be equal across countries while allowing short-run dynamics and error-correction speeds π_i to vary freely. It occupies the natural middle ground between homogeneous FE and fully heterogeneous MG estimators: more efficient when pooling is valid, and the Hausman test determines validity. Table 7 reports full results for the *quantity* specifications.

Both Hausman tests are passed with comfortable margins (AMG vs PMG: $H = 3.16$, $p = 0.367$; CCEMG vs PMG+CSD: $H = 3.07$, $p = 0.381$), formally confirming that long-run homogeneity is valid: the cross-country distribution of estimated long-run KPF coefficients is tight enough that pooling is both valid and more efficient. The PMG+CSD error-correction speeds, 77% of which are negative (stabilising), with median $\hat{\pi}_i = -0.457$ (implied half-life of 1.5 years), are consistent with the cointegrating interpretation and indicate that deviations from the long-run knowledge production equilibrium are corrected rapidly.

Steady-state calibration. A natural theoretical check is whether the estimated parameters imply a balanced-growth-path (BGP) rate of scientific output growth consistent with what we observe in the data. On the BGP, the long-run equilibrium path along which knowledge output, human capital, and the knowledge stock all grow at constant rates, science output grows at $g_A^* = \sigma \cdot n_H/(1-\phi)$, where n_H is the growth rate of the research labour force. The relevant observed benchmark is the above-median sample mean publication growth rate of approximately 5%–6%/yr over 2003–2022.

Using the preferred PMG+CSD quantity estimates ($\hat{\sigma} = 1.825$, $\hat{\phi} = 0.652$) with $n_H = 0.66\%$/yr (the panel-mean PWT human capital growth rate for the above-median sample), the implied BGP rate is $g_A^* = 1.825 \times 0.0066/(1 - 0.652) \approx 3.5\%$/yr. The AMG Spec B estimate ($\hat{\sigma} = 2.30$, $\hat{\phi} = 0.762$) implies $g_A^* \approx 6.4\%$/yr, and the FE Spec B estimate ($\hat{\sigma} = 0.166$, $\hat{\phi} = 0.646$) implies $g_A^* \approx 0.3\%$/yr. Comparing the three implied BGP rates against this benchmark: the PMG+CSD-implied rate (3.5%/yr) lies below the sample mean, the AMG-implied rate (6.4%/yr) lies above it, and the FE-implied rate (0.3%/yr) lies far below. The PMG+CSD and AMG estimates therefore straddle the observed mean, which is itself consistent with the above-median sample being in transitional dynamics rather than at a BGP steady state during 2003–2022. The rapid expansion of China, India, and Brazil was pulling the observed growth rate above the steady-state level implied by the panel-mean parameter estimates. The FE estimate is

clearly at odds with the observed growth experience and reinforces the interpretation that FE under the PWT hc specification severely attenuates $\hat{\sigma}$ due to the smooth within-country variation of the human capital index.

National versus global knowledge stock: the rest-of-world robustness test. Section 2 introduced the theoretical question of whether knowledge production is governed by a country's own accumulated science base or by the global knowledge pool. The augmented specification (5) provides a direct empirical test; full results are in Appendix D, Table D.1. Three findings stand out (Table 10). First, under FE, the only estimator with clean identification in this specification (Section 4.2), the own-stock coefficient is entirely stable: $\hat{\phi}_1 = 0.645^{***}$ vs 0.646^{***} in the baseline, a change of 0.001. Own-stock compounding is a genuine domestic effect, not a proxy for global frontier proximity. Second, the FE estimate of the rest-of-world spillover is $\hat{\phi}_2 = 0.326$ (SE = 0.336, $p \approx 0.33$): positive and directionally consistent with Coe–Helpman spillovers, but imprecise, reflecting the thin cross-sectional identifying variation in $\ln K_{t-1}^{-i}$ (within-year SD = 0.036). Third, the collapse of $\hat{\phi}_1$ in AMG and the explosion of $\hat{\phi}_2$ in CCEMG confirm the identification derivation of Appendix D: these estimators were already absorbing the global frontier through their internal controls, so adding $\ln K_{t-1}^{-i}$ induces near-collinearity with those controls. The instability is confirmatory, not alarming.[21]

Table 10: Rest-of-World Robustness: Baseline vs Augmented Specification Summary (Spec B, Above-Median)

	Baseline (own K only)			Augmented (own + RoW K)		
Estimator	$\hat{\sigma}$	$\hat{\gamma}$	$\hat{\phi}$	$\hat{\phi}_1$	$\hat{\phi}_2$	Verdict
FE	0.214	0.046^{**}	0.646^{***}	0.645^{***}	0.326	$\hat{\phi}_1$ stable
AMG	2.927^{***}	0.076	0.715^{***}	0.195	0.808^{**}	CDP absorbs RoW trend
CCEMG	1.150	−0.007	0.667^{***}	0.049	$27.18^{***\dagger}$	CS-means absorb RoW

Notes: $\hat{\phi}_1$ = own-stock elasticity; $\hat{\phi}_2$ = rest-of-world spillover. †CCEMG $\hat{\phi}_2$ is a near-singular numerical artefact. FE is the only estimator with clean identification of both coefficients. Full results in Appendix D, Table D.1. ***, **: significance at 1%, 5%.

R&D expenditure and the knowledge stock: robustness check. Table 11 addresses the concern that $\hat{\phi}$ proxies for sustained national R&D investment rather than genuine knowledge-stock compounding. Baseline columns reproduce the Spec B results

[21] A natural question is whether PMG could be applied to the augmented rest-of-world specification. The answer is no, for the same structural reason that prevents AMG and CCEMG: PMG's long-run pooling equation embeds the world knowledge level in its pooled coefficients on $\ln K_{i,t-1}$; augmenting with $\ln K_{t-1}^{-i}$ would introduce near-collinearity with both the PMG long-run equation and the CSD augmenter in PMG+CSD. The identification argument of Appendix D applies to PMG as directly as to AMG and CCEMG: the global knowledge level is absorbed by each estimator's internal controls, leaving only the large-economy gradient to identify $\hat{\phi}_2$ under FE.

Table 11: Robustness Check 6: R&D Expenditure Augmented KPF Specifications

	Baseline	$+\ln$ GERD
Panel A: Quantity — Fixed Effects		
$\hat{\sigma}$: $\ln H_A$	0.166	−0.150
	(0.656)	(0.460)
$\hat{\gamma}$: $\ln Q$	0.046	−0.042
	(0.045)	(0.047)
$\hat{\phi}$: $\ln K_{t-1}$	0.646***	0.637***
	(0.014)	(0.006)
$\ln$ GERD	—	0.232**
		(0.097)
Obs. / Countries	1,383 / 81	1,106 / 77
Panel B: Quantity — AMG		
$\hat{\sigma}$: $\ln H_A$	2.300***	1.905**
	(0.802)	(0.768)
$\hat{\gamma}$: $\ln Q$	−0.048	0.024
	(0.103)	(0.100)
$\hat{\phi}$: $\ln K_{t-1}$	0.762***	0.721***
	(0.116)	(0.092)
$\ln$ GERD	—	0.056*
		(0.032)
Obs. / Countries MG	1,383 / 73	1,106 / 68
Panel C: Quantity — CCEMG		
$\hat{\sigma}$: $\ln H_A$	1.056	0.946
	(0.910)	(0.876)
$\hat{\gamma}$: $\ln Q$	−0.011	−0.019
	(0.098)	(0.124)
$\hat{\phi}$: $\ln K_{t-1}$	0.751***	0.700***
	(0.107)	(0.128)
$\ln$ GERD	—	0.039
		(0.036)
Obs. / Countries MG	1,383 / 68	1,106 / 68
Panel D: Quality — PMG+CSD (the valid estimator for this specification)		
$\hat{\sigma}_q$: $\ln H_A$	2.188***	1.464***
	(0.196)	(0.203)
$\hat{\gamma}_q$: $\ln Q$	0.039**	−0.006
	(0.018)	(0.016)
$\hat{\phi}_q$: $\ln K_{t-1}$	0.015**	0.009
	(0.007)	(0.006)
$\ln$ GERD	—	0.038*
		(0.021)
Obs. / Countries	1,527 / 77	1,163 / 60

Notes: Panels A–C: dependent variable ln(whole-count articles), above-median sample. Baseline columns reproduce Spec B results from Tables 4–5; $+\ln$ GERD columns estimated on the GERD-available sub-sample. Panel D: logit-transformed top-1% citation share; PMG+CSD is the preferred estimator for the quality specification given the I(0)/I(1) mismatch (see Section 4.2); baseline replicates Table 12 exactly. $\ln \text{GERD}_{it} = \ln(g_{it} \times Y_{it})$; g_{it} = GERD/GDP from WDI (World Bank, 2026); Y_{it} = real GDP from PWT 11.0 (Feenstra et al., 2015); covers 77 of 81 above-median countries (Panels A–C), 60 countries (Panel D). Entity-clustered standard errors in parentheses (Panels A–C). ***, **, *: 1%, 5%, 10%.

from Tables 4–5; + ln GERD columns are estimated on the GERD-available sub-sample (77 countries, $N = 1{,}106$).

Two findings stand out. *First*, $\hat{\phi}$ is essentially unchanged after conditioning on GERD across all three estimators: FE moves from 0.646^{***} to 0.637^{***} (-1%), AMG from 0.762^{***} to 0.721^{***} (-5%), and CCEMG from 0.751^{***} to 0.700^{***} (-7%). None of these movements is economically meaningful. The knowledge stock effect is not a proxy for R&D expenditure intensity. *Second*, the GERD coefficient divides along estimator lines in a pattern consistent with the identification logic: GERD is significant in FE (0.232^{**}, SE = 0.097) and marginally significant in AMG (0.056^{*}, SE = 0.032), but entirely insignificant in CCEMG (0.039, SE = 0.036). FE captures cross-country *level* differences in R&D intensity not accounted for elsewhere; CCEMG, whose cross-sectional mean augmenters absorb the global frontier non-parametrically, finds no residual within-country time-series GERD effect. The AMG result (marginally significant) likely reflects partial redundancy between AMG's Common Dynamic Process and GERD, both of which track global knowledge-accumulation trends. The modest reduction in $\hat{\sigma}$ under FE reflects multicollinearity between human capital quality and R&D intensity, which inflates standard errors without displacing $\hat{\phi}$.

5.3 Discussion: The Quality Specification - Citation Impact

Quantity versus quality. The quality specifications (Table 12) reveal a structurally distinct production process. Under PMG+CSD — the preferred estimator for this specification, for reasons given below — human capital is the principal driver of citation impact: $\hat{\sigma}_q = 2.188$ ($t = 11.16$), with $\hat{\gamma}_q = 0.039$ ($t = 2.21$) and $\hat{\phi}_q = 0.015$ ($t = 2.09$). The knowledge stock elasticity is two orders of magnitude below its quantity counterpart ($\hat{\phi} \approx 0.65$): accumulated knowledge confers no compounding advantage for citation impact, only for volume. Human capital quality, not the size of the existing knowledge base, determines whether a country produces frontier-cited science.

One further results from PMG+CSD is notable: the positive and significant $\hat{\gamma}_q = 0.039$ ($t = 2.21$) is consistent with the complementarity interpretation flagged in Section 3: a denser commercial knowledge environment raises the probability of producing frontier-cited output.

The estimator choice requires brief explanation. The logit citation share is I(0) while the regressors are I(1) (Table 3): there is no cointegrating relationship between the dependent variable and the regressors by construction, so FE, AMG, CCEMG, and standard PMG — all designed for I(1)/I(1) systems or long-run cointegrating relationships — are not well suited to this specification. Their $\hat{\sigma}_q$ estimates confirm this: the range runs from -10.30 (AMG) to $+1.26$ (FE), an empirical signature of estimators applied outside their designed domain. PMG+CSD absorbs cross-sectional dependence through cross-

sectional mean augmenters in the short-run equation without requiring a cointegrating structure, removing the specific source of misspecification. No formal consistency proof exists for PMG+CSD in the I(0)/I(1) case; we advance it as the preferred estimator on methodological and empirical grounds, noting that its result is stable across all reasonable perturbations of the specification (Section 4.2).

One further result from the quality estimation is notable. The FE specification corroborates $\hat{\phi}_q \approx 0$ through a different channel: FE does not absorb the global frontier, so if the knowledge stock independently raised citation quality beyond what the frontier already provides, $\hat{\phi}_q$ would appear under FE. It does not. The near-zero knowledge stock elasticity is therefore not an artefact of frontier absorption by the CSD-corrected estimators, but a genuine structural feature of the quality production process.

This result is complementary to, not a substitute for, the country fixed-effects analysis below. The two findings answer different questions. PMG+CSD identifies the structural elasticities: how much citation quality responds to a unit change in human capital, conditional on the global frontier. By contrast the fixed-effects decomposition asks how much of the *level* of cross-country variation in citation quality is explained by time-invariant institutional characteristics versus the measured inputs. Both can be true simultaneously: human capital quality can be a strong and well-identified *driver* of citation impact (the PMG+CSD slope), while a large share of the *level* of cross-country variation is nonetheless explained by a time-invariant country characteristic (the fixed-effect decomposition) if, as is plausible, the institutional characteristic itself partly operates by shaping how effectively a country's human capital is deployed and rewarded internationally. Table 13 reports $\hat{\mu}_{i,q}$ for fourteen of the fifteen focus economies; Türkiye is excluded because the logit citation share is missing for a sufficient number of years to preclude a stable FE estimate in the above-median sample.

Country fixed effects in the quality specification. The same FE extraction procedure used to construct Table 9 for quantity, applied to the quality specification (Spec Q-B), provides a complementary decomposition to the PMG+CSD slope estimates above. Again, this decomposition uses the FE estimator, which is not the preferred estimator for recovering the *structural elasticities* of the quality specification, given the I(0)/I(1) mismatch (see Section 5.3); the fixed-effects decomposition is used here for a different purpose (characterising the cross-country *level* of variation explained by time-invariant institutional characteristics versus measured inputs) for which the FE within-country demeaning is an appropriate and standard tool independent of the cointegration concerns that affect the slope coefficients. The decomposition shows that the country fixed effect $\hat{\mu}_{i,q}$ accounts for 65% of the cross-country variance in citation quality, while the slope coefficients in the FE regression jointly account for only 8%, a near-complete inversion of the quantity specification, where slopes explain 63% of cross-country variance and the fixed

Table 12: Quality Specification – Logit-Transformed Top-1% Citation Share

	Spec Q-A: H = Researchers			Spec Q-B: H = PWT hc				
	FE	AMG	CCEMG	FE	AMG	CCEMG	PMG	PMG+CSD†
$\hat{\sigma}_q$: $\ln(H_A)$	0.240***	0.032	0.059	1.260**	−10.297**	−1.131	−0.393	2.188***
	(0.075)	(0.214)	(0.215)	(0.534)	(4.057)	(4.181)	(0.386)	(0.196)
$\hat{\gamma}_q$: $\ln(Q)$	−0.029	−0.115	−0.091	0.007	0.456	0.276	0.377***	0.039**
	(0.046)	(0.550)	(0.431)	(0.035)	(0.404)	(0.369)	(0.035)	(0.018)
$\hat{\phi}_q$: $\ln(K_{t-1})$	0.012	0.899	0.073	0.016	0.226	−0.117	−0.040***	0.015**
	(0.015)	(1.023)	(0.322)	(0.016)	(0.370)	(0.444)	(0.014)	(0.007)
Countries in MG	80	53	49	81	72	66	77	77
Observations		914				1,232		

Notes: Dependent variable: $\text{logit}(p_{it})$, the logit-transformed share of country i's publications in the global top-1% citation percentile in year t. Above-median sample only. Standard errors in parentheses. ***, **: significance at 1%, 5%. PMG = Pooled Mean Group (Pesaran et al., 1999); PMG+CSD = PMG augmented with cross-sectional means of all variables as short-run regressors.

†**PMG+CSD is the preferred estimator for the quality specification.** The logit citation share is I(0) while $\ln(H_A)$, $\ln(Q)$, and $\ln(K_{t-1})$ are I(1) (Table 3): FE, AMG, CCEMG, and standard PMG are not designed for this integration-order mismatch, which explains the instability in $\hat{\sigma}_q$ across those estimators (range: -10.3 to $+1.3$; standard PMG: -0.393, $t = -1.02$). PMG+CSD absorbs cross-sectional dependence via cross-sectional mean augmenters in the short-run equation without requiring a cointegrating structure, and is the only estimator to recover a stable, precisely estimated human capital coefficient: $\hat{\sigma}_q = 2.188$ ($t = 11.16$). See Section 5.3.

effect only 5%. This finding is fully compatible with the PMG+CSD result that human capital quality is a strong structural driver of citation impact: the FE slopes are weak not because the underlying relationship is absent (PMG+CSD shows it is not) but because FE's within-country demeaning, combined with the I(0)/I(1) mismatch, leaves little usable variation for the slope coefficients to estimate from once a large, persistent country effect is partialled out. The substantive reading is therefore additive rather than competing: a country's citation impact is shaped both by its human capital quality (the PMG+CSD structural elasticity) and by a separate, time-invariant institutional or scientific-culture component (the $\hat{\mu}_{i,q}$ level effect) that plausibly governs how effectively that human capital is deployed internationally, for instance through research culture, international collaboration norms, or the domestic concentration of frontier-level researchers. Table 13 reports $\hat{\mu}_{i,q}$ and the corresponding intercept for the same focus economies. Because the dependent variable is the logit-transformed top-1% citation share, the grand-mean intercept itself is large and negative (-6.629): this is an artefact of the logit transformation of a small baseline probability and carries no substantive interpretation on its own. What matters is the *relative* position of each country's $\hat{\mu}_{i,q}$ and the implied intercept, exactly as for the quantity table.

Three findings follow. First, the ranking inverts substantially relative to quantity: the rank correlation between $\hat{\mu}_i$ and $\hat{\mu}_{i,q}$ across the 81-country mean-group sample is only 0.47, confirming these are empirically distinct institutional dimensions rather than two measures of the same underlying NIS quality. Second, the United States and United Kingdom are the only focus economies that rank highly on *both* quantity and quality fixed effects, consistent with mature, internationally networked research systems that combine scale with citation impact. Third, and most directly relevant to the projections in Section 5.4.1, China's citation-quality fixed effect is unremarkable (rank 40 of 81) despite its dominant quantity fixed effect (rank 1 of 81): China's path to citation-quality leadership, if it occurs, cannot be read off its current quantity-side institutional advantage, and would instead require closing the *quality-specific* institutional gap captured by $\hat{\mu}_{i,q}$, a process the Japan and South Korea precedents (Online Appendix A.4) suggest typically lags volume catch-up by a decade or more.

R&D expenditure and citation quality: robustness check. Panel D of Table 11 reports the quality specification augmented with $\ln \text{GERD}$, estimated by PMG+CSD — the most justifiable estimator for this specification given the I(0)/I(1) mismatch (see Section 5.3). The baseline replicates Table 12 exactly ($\hat{\sigma}_q = 2.188^{***}$, $\hat{\gamma}_q = 0.039^{**}$, $\hat{\phi}_q = 0.015^{**}$). Adding $\ln \text{GERD}$ produces three findings. First, GERD itself is borderline significant (0.038^{*}, SE $= 0.021$, $t = 1.79$), though this should be read with caution: the sample shrinks from 77 to 60 countries when GERD coverage is required, and the marginal significance is partly a sample-composition effect. Second, $\hat{\phi}_q$ falls from 0.015^{**}

Table 13: Country Fixed Effects ($\hat{\mu}_{i,q}$) in the Quality Specification, FE Spec Q-B (Above-Median)

ISO	Country	$\hat{\mu}_{i,q}$	Intercept†	Rank (Quality)
USA	United States	+0.894	−5.735	4
GBR	United Kingdom	+0.795	−5.834	8
CAN	Canada	+0.618	−6.011	16
FRA	France	+0.489	−6.140	25
ITA	Italy	+0.478	−6.151	26
DEU	Germany	+0.440	−6.189	29
IND	India	+0.402	−6.227	30
ISR	Israel	+0.335	−6.293	35
ZAF	South Africa	+0.353	−6.276	34
CHN	China	+0.174	−6.455	40
KOR	South Korea	−0.164	−6.793	49
BRA	Brazil	−0.304	−6.933	57
JPN	Japan	−0.333	−6.962	59
MEX	Mexico	−0.446	−7.075	64
Panel reference				
–	Panel mean (81 countries)	0.000	−6.629	
–	Panel SD	0.711	–	

Notes: $\hat{\mu}_{i,q}$ estimated from two-way FE Spec Q-B (logit top-1% citation share, PWT hc index, above-median sample, $N = 81$), normalised to sum to zero across the mean-group sample. †Intercept = grand-mean intercept (−6.629) $+\hat{\mu}_{i,q}$. **Negative values in this table mean something different from the quantity table and require separate explanation.** Because the dependent variable is a logit-transformed citation *share*, the natural reference point is zero ($\hat{\mu}_{i,q} = 0$), corresponding to the panel average; values are therefore best read purely as a relative ranking, not against an absolute output benchmark. A negative $\hat{\mu}_{i,q}$ indicates a country's logit citation share is below the panel-average *conditional* relationship — for example, Japan (−0.33) and Mexico (−0.45) both have lower-than-predicted citation quality given their inputs, despite Japan recording a strongly *positive* fixed effect in the quantity specification ($\hat{\mu}_i = +0.27$, rank 19; Table 9). These two facts are compatible: Japan's research system efficiently converts inputs into article *volume* but underperforms the panel-average citation-quality relationship, illustrating precisely the structural asymmetry between quantity and quality compounding documented in this section. China's modest positive $\hat{\mu}_{i,q}$ (+0.17, rank 40 of 81 — essentially mid-panel) is the most consequential entry for the paper's forward projections: it confirms that China's enormous quantity fixed effect ($\hat{\mu}_i = +1.26$, rank 1) does not carry over to citation quality, where China sits near the panel median.

to 0.009 (SE = 0.006) and loses significance; an expected consequence of multicollinearity between the knowledge stock and R&D expenditure in the smaller sample, not evidence of a structural link. Third, and most importantly, $\hat{\sigma}_q$ remains large and highly significant at 1.464^{***} (SE = 0.203, $t = 7.22$): human capital quality is the principal driver of citation impact even after conditioning on R&D investment. The structural asymmetry, quality driven by human capital, not knowledge-stock compounding, is fully robust to inclusion of R&D expenditure.

5.4 Scale Dynamics and Projected Knowledge Production Shares, 2022–2047

The results of Sections 5.2 and 5.3 characterise how inputs into the KPF are converted to outputs, in volume and quality metrics respectively. A complementary question concerns *scale dynamics*: how will the global distribution of knowledge production shift over the coming decades if countries continue to expand their inputs at the rates observed during the sample period? This section provides a quantitative answer by simulating the KPF forward 25 years to 2047, considering the implications for both the volume (Section 5.4.1) and quality (Section 5.4.2) of knowledge production.

We state a core caveat regarding the projections which follow up front: 20 year projections are ambitious under the very best of circumstances, and here this is compounded by the fact that the data set has only 20 observations per country. Inevitably, confidence intervals around any projections undertaken as formal forecasts are large, and this is before one considers the possibility of demographic changes, changes in science-related and other institutions, as well as other structural breaks. It is therefore imperative to understand that the exercise presented here is one of noting the result of the point estimate consequences that will follow in compound growth terms if scientific knowledge production functions remain stable, and countries continue with their current rate of investment in the inputs into knowledge production. As such, they are best understood as a scenario projection: they are not intended to claim forecast accuracy, and should not be overinterpreted.

5.4.1 Volume of Knowledge Production

Projection design. The projection uses the panel mean-group KPF elasticities (**AMG, above-median sample:** $\hat{\sigma} = 2.30$, $\hat{\gamma} = -0.05$, $\hat{\phi} = 0.76$) for all countries. By imposing a common production technology, differences in projected shares arise solely from differences in initial conditions (2022 knowledge stocks and output levels) and in the rate at which countries accumulate inputs, not from differences in KPF efficiency. Country-specific intercepts are calibrated so that each country's projected 2022 output matches its actual 2022 publication count exactly.

For each country, human capital (H) grows at its estimated in-sample log-linear trend rate ($\bar{g}_H$, 2003–2022). Patent stock growth rates (g_Q) are similarly estimated from the in-sample trend but capped at the 75th percentile across countries (9.83% per annum) to moderate the mechanical extrapolation of unsustainably rapid patent expansions recorded in some economies during the sample period. The knowledge stock (A) is updated endogenously each period via the perpetual inventory recursion $A_{i,t} = (1 - \rho_K)\, A_{i,t-1} + \dot{A}_{i,t}$ with $\rho_K = 0.10$, so that projected publication output in each period feeds back into the knowledge stock available in the next, thereby preserving the compounding mechanism documented in Sections 5.2.

Table 14 reports the input trend growth rate assumptions for fifteen focus countries. Table 15 presents projected shares of above-median world knowledge output at five-year intervals over 2022–2047. Figure 1 provides the full visual summary.

Caveats. These projections are conditional on three assumptions that deserve explicit acknowledgement. First, input growth rates are extrapolated from the 2003–2022 sample at constant log-linear rates; in reality, human capital growth rates are likely to decelerate as science systems approach developed-country educational attainment levels. Second, the projection assumes structural stability of the KPF across the horizon, including the compounding parameter $\hat{\phi}$; if diminishing returns set in at larger knowledge stock levels, as the Japan finding suggests, the projected gains for rapidly growing economies are overstated. Third, the above-median sample covers 77 countries, and the 'rest of sample' share grows from 33.6% to 43.7% over the horizon, reflecting the expansion of science systems not individually analysed here. Notwithstanding these qualifications, the projections provide a disciplined, model-consistent basis for assessing the scale dynamics of global knowledge production, a complement to the efficiency analysis of the preceding sections.

Core findings are enumerated below.

China: peak share in the mid-2030s, then partial retrenchment. The projections assign China the largest gain in global share among the fifteen focus countries, rising from 22.6% in 2022 to a peak of approximately 29.8% around 2040, before retreating modestly to 27.8% by 2047. The rise reflects China's combination of a large and still-growing knowledge stock, with the largest absolute compounding base in the sample at the common elasticity $\hat{\phi} = 0.76$, and its patent stock growing at the cap rate of 9.83% per annum. (The projection uses panel-mean elasticities for all countries; China's large gain reflects its large initial stock and fast input growth, not a country-specific efficiency advantage.) The mild retrenchment after 2040 reflects the diminishing marginal return to the patent stock ($\hat{\gamma} < 0$ in the panel specification): as China's patent stock reaches very large absolute levels, the slightly negative elasticity moderates its contribution to output

Table 14: Input Trend Growth Rate Assumptions for the Projection

ISO	Country	$\bar{g}_H$ (%/yr, HC)	$\bar{g}_Q$ (%/yr, patent stock, capped)
CHN	China	0.87	9.83
IND	India	1.08	9.83
TUR	Türkiye	1.28	9.83
MEX	Mexico	0.71	6.71
ITA	Italy	0.66	4.80
BRA	Brazil	2.28	3.64
CAN	Canada	0.27	3.73
USA	United States	0.23	4.11
FRA	France	0.60	2.99
GBR	United Kingdom	0.33	1.25
DEU	Germany	0.13	2.31
JPN	Japan	0.35	0.04
ZAF	South Africa	1.19	−0.08
ISR	Israel	0.86	6.30
KOR	South Korea	0.81	5.79
Panel median	(77 countries)	0.66	4.82

Notes: $\bar{g}_H$ and $\bar{g}_Q$ are OLS log-linear trend slopes estimated over 2003–2022. Patent stock growth rates capped at the 75th percentile across all 77 countries (9.83%/yr) to moderate extrapolation of historically rapid patent expansions. Human capital index from PWT 11.0 (Feenstra et al., 2015); patent stock from WIPO (World Intellectual Property Organization, 2026), constructed via PIM with $\rho_P = 0.15$.

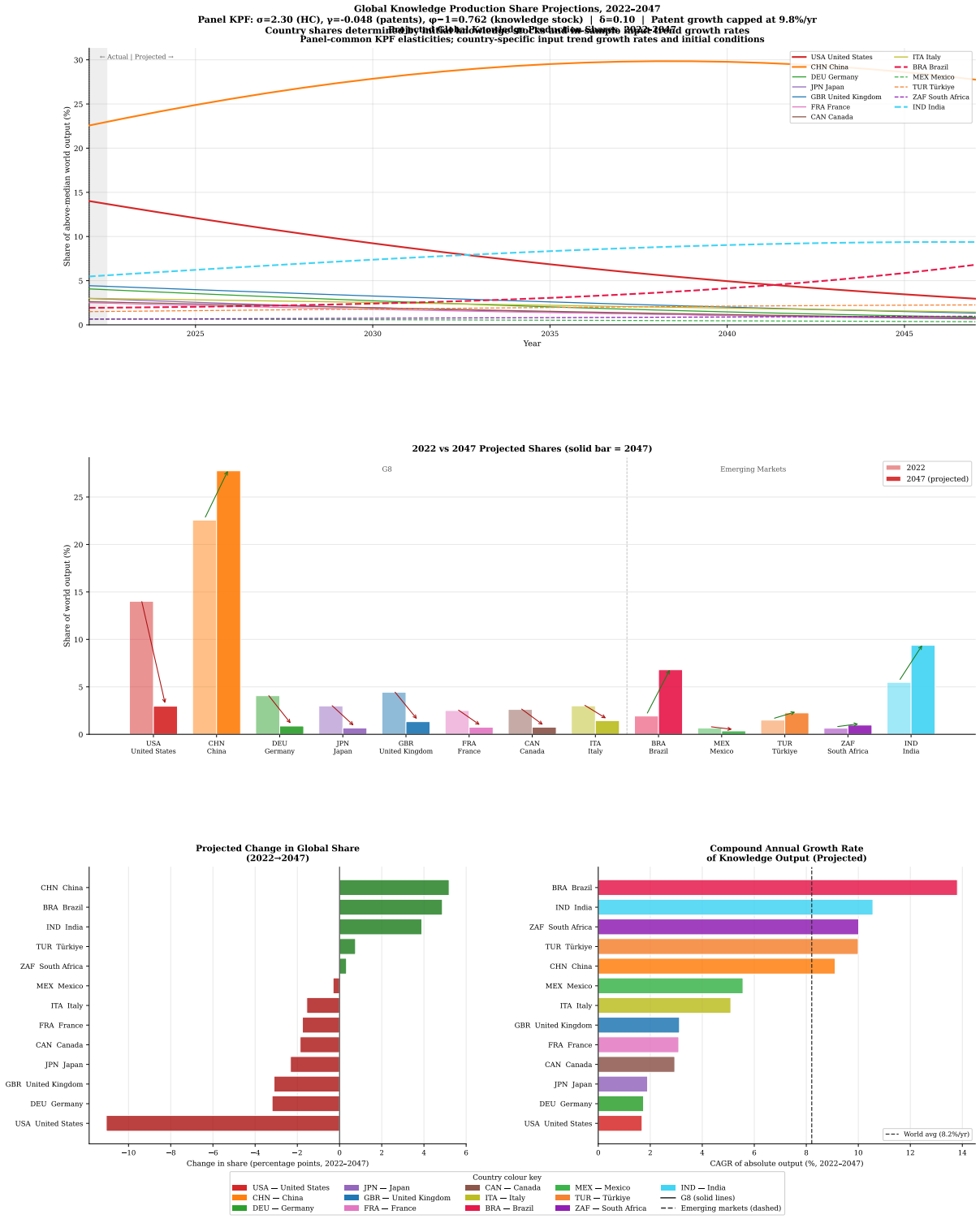


Figure 1: Projected Global Knowledge Production Shares, 2022–2047.
Notes: Panel-common KPF elasticities applied to all countries; country-specific initial conditions and input trend growth rates. Top panel: time series of projected shares. Middle panel: 2022 vs 2047 bar comparison; arrows indicate direction of change. Bottom-left: projected percentage-point change in share (green = gain, red = loss). Bottom-right: compound annual growth rate of absolute knowledge output; dashed line = world average (8.2%/yr). G8 economies shown with solid lines; emerging markets with dashed lines.

Table 15: Projected Shares of Global Knowledge Output (%), 2022–2047

ISO	Country	2022	2025	2030	2035	2040	2047
CHN	China	22.57	24.89	27.86	29.52	29.78	27.76
IND	India	5.48	6.22	7.37	8.34	9.03	9.38
BRA	Brazil	1.93	2.04	2.41	3.05	4.13	6.80
USA	United States	14.01	12.09	9.23	6.85	4.94	2.96
TUR	Türkiye	1.50	1.61	1.79	1.96	2.11	2.25
ITA	Italy	3.00	2.83	2.53	2.21	1.89	1.45
GBR	United Kingdom	4.43	3.98	3.26	2.60	2.01	1.33
ZAF	South Africa	0.64	0.68	0.75	0.82	0.89	0.97
ISR	Israel	0.54	0.50	0.45	0.40	0.34	0.28
KOR	South Korea	2.22	2.13	1.96	1.78	1.59	1.29
DEU	Germany	4.07	3.54	2.73	2.03	1.47	0.87
CAN	Canada	2.62	2.36	1.94	1.53	1.17	0.75
FRA	France	2.50	2.20	1.75	1.39	1.09	0.75
JPN	Japan	2.99	2.55	1.94	1.45	1.06	0.67
MEX	Mexico	0.66	0.64	0.58	0.52	0.46	0.36
Rest of sample	33.61	34.37	35.87	37.72	39.97	43.71	

Notes: Shares of above-median sample output (77 countries). Panel-common KPF elasticities: $\hat{\sigma} = 2.30$, $\hat{\gamma} = -0.05$, $\hat{\phi} = 0.76$. Country intercepts calibrated to match actual 2022 output. Knowledge stock updated endogenously via PIM, $\rho_K = 0.10$. G8 economies above the dividing rule; emerging markets below.

growth. By 2047, China accounts for more than a quarter of all knowledge output among science-active economies in our sample, up from under a quarter today.

The United States: structural decline in share. The most striking projected shift is the sustained compression of the United States' share, from 14.0% in 2022 to 2.96% by 2047, a loss of 11 percentage points, the largest decline of any country in the group. This reflects the combination of two factors that are equally structural in character. First, the US human capital index grows at only 0.23% per annum, the second lowest in the group, potentially reflecting a largely saturated high-quality research workforce. Second, the US patent stock grows at 4.1% per annum, a moderate rate that generates no scale advantage. Crucially, the US knowledge stock, although large in absolute terms and still growing, cannot compensate for slow input expansion when the rest of the world is expanding faster, because the common $\hat{\phi} = 0.76$ is identical across countries. The implication is stark: US dominance of global knowledge production in 2047 depends on whether it can accelerate input accumulation above its recent trend. To maintain its current 14% global share through 2047, the United States would need to produce approximately 5.3 times as much knowledge output in 2047 as the baseline projection delivers. Under the KPF framework with $\hat{\sigma} = 2.30$, this translates to a required HC quality growth rate of approximately 3.0%/yr; roughly 13 times the current rate of 0.23%/yr and approximately

5 times the exceptional post-WWII US expansion rate ($\approx$ 0.5–0.7%/yr during the GI Bill era). No plausible acceleration of human capital investment along historical precedents would be sufficient to maintain the current US share. The structural trajectory is driven by the compounding advantage of rapidly growing knowledge stocks elsewhere, not merely by US policy stances.

Emerging market gains and G8 retreat. Among emerging markets, Brazil records the largest gain (+4.9 pp, driven by the fastest HC growth in the group at 2.3%/yr), India approaches 9.4% of world output, and Türkiye grows steadily on strong HC and patent stock expansion. Mexico illustrates the efficiency-versus-scale paradox: despite the highest *point-estimate* HC elasticity in the analysis ($\hat{\sigma}_i \approx 4.4$ on average across combinations, though with wide uncertainty), it *loses* share because its knowledge stock is too thin to generate compounding returns and its HC grows at only 0.71%/yr, consistent with significant brain-drain to the United States. Among the G8, every non-China member loses share, ranging from modest (Italy, Canada, France: -1.6 to -1.9 pp) to severe (Japan -2.3 pp, UK -3.1 pp, Germany -3.2 pp). Japan's share collapses to 0.67% by 2047, consistent with its near-zero patent stock growth and negative knowledge stock elasticity. The projections use panel-mean elasticities throughout; country-specific projections with trimmed AMG coefficients exhibit extreme sensitivity at $T = 20$, validating the panel-mean approach.

Limitations and the case for projection. Two observations balance the technical caveats that follow. These projections describe the consequences of the *status quo*, of carrying on with current input accumulation trajectories. Ignoring such projections risks policy malpractice: historically, every displacement of a scientific leader (England by France, France by Germany, Germany by the United States) was preceded by decades during which the incumbent underestimated the pace of catch-up. The projected eleven-percentage-point decline for the United States is not a forecast of imminent collapse but a structural trajectory that is already underway (low human capital growth in the US, high elsewhere) and that compounds through the $\hat{\phi}$ mechanism. Note also that quantity and quality projections tell different stories: the quality-weighted analysis (Section 5.4.2) reveals a more modulated US decline.

Convergence sensitivity. As a robustness check, we examine two scenarios in which each country's input growth rates converge linearly or exponentially to the G5 frontier average ($g_H^* = 0.33\%$/yr, $g_Q^* = 2.11\%$/yr) by 2047. Full results are in Appendix C.

Six country-specific findings are worth flagging across baseline and convergence scenarios. *China* gains more under convergence than in the baseline (its 2047 share rises toward 30% under exponential convergence) because the common KPF elasticity applied

to China's large initial knowledge stock generates more additional output than the same elasticity applied to smaller-stock countries, even when all countries' HC growth rates converge to the same frontier average; convergence does not equalise outcomes when initial conditions are highly unequal. *The United States* declines less steeply under convergence (-9 pp under exponential convergence vs -11 pp in the baseline): convergence raises the US HC growth rate from its very low baseline (0.29%/yr) toward the frontier average (0.33%/yr), partially offsetting the decline, though the offset is modest because no plausible HC acceleration is sufficient to offset the compounding advantage of rapidly growing knowledge stocks elsewhere. *India* and *Türkiye*, whose in-sample HC growth rates (1.02% and 1.23%/yr respectively) already exceed the frontier target, see their HC growth rates pulled *down* under convergence; their gains relative to the baseline therefore reflect the patent stock convergence channel, which adds to their already-positive patent accumulation rates. *Brazil* is the most scenario-sensitive country because its exceptional HC growth rate (2.28%/yr) is reduced most sharply by convergence, making its projected share highly sensitive to the convergence speed assumed: its 2047 share ranges from approximately 4% (exponential convergence, which curtails its HC advantage most aggressively) to above 9% (baseline, where its full HC growth rate is maintained). *Japan* reaches near-zero share under all specifications, confirming that its structural decline is driven by near-zero patent stock growth and negligible HC accumulation rather than by any assumption about convergence trajectories. The headline result, a multipolar science world by 2047 with China accounting for roughly a quarter to nearly a third of world output, holds across all scenarios.

5.4.2 Quality-Adjusted Projections: Who Will Lead in High-Impact Science?

A structurally distinct question concerns *high-impact* science: which countries would our findings suggest will account for the largest shares of globally top-cited research? We address this using the PMG+CSD quality KPF estimates, which Section 5.3 establishes as the methodologically preferred estimator for the quality specification.

Approach A (primary): direct quality projection from factor inputs. The primary approach projects each country's top-1% citation share directly from factor input growth via the quality KPF alone, without reference to the volume projection. Country-specific intercepts α_i^q are calibrated to match each country's 2020 top-1% citation fraction (the last year with sufficient citation accumulation), and the projection evolves the logit citation share forward using in-sample HC and patent stock growth rates:

$$\text{logit}(f_{it}^{top1}) = \alpha_i^q + \hat{\sigma}_q \cdot \ln H_{it} + \hat{\gamma}_q \cdot \ln Q_{it} + \hat{\phi}_q \cdot \ln K_{i,t-1}, \tag{17}$$

with $\hat{\sigma}_q = 2.188$, $\hat{\gamma}_q = 0.039$, $\hat{\phi}_q = 0.015$ (PMG+CSD, Spec Q-B). Each country's share of world top-cited output is then the share of its projected top-1% article count in the cross-country total. This approach asks: based solely on how human capital quality and other factor inputs evolve, what happens to each country's share of globally top-cited science?

The central finding is that *the sharp G8 volume attenuation is substantially arrested.* Table 16 (Approach A columns) shows the United States retaining a 17.1% share of world top-cited science by 2047, compared to 21.5% in 2022: a decline of only −4.4 percentage points, in sharp contrast to the −11.0 pp volume loss. The United Kingdom falls from 7.4% to 6.2% (−1.2 pp); Germany from 4.1% to 3.1% (−1.0 pp); France retains almost its full 2022 share (2.54% to 2.46%). The contrast with the volume projections is stark: quality leadership, driven by human capital dynamics rather than knowledge-stock compounding, declines far more slowly than volume leadership. China's top-1% share rises from 21.2% to 23.8% (+2.7 pp), a meaningful gain, but far more modest than the volume gain or than Approach B suggests. Among emerging markets, India (3.1% to 3.9%), Brazil (0.9% to 2.2%), and Turkey (1.0% to 1.3%) all gain, reflecting their positive HC growth rates.

Limitation and Approach B (qualifier): incorporating volume dynamics. Approach A holds the article base constant, projecting quality purely from factor input dynamics. While isolating projections from the quality KPF, this does not account for a potential impact from the growth in total article volume. A country that doubles its article output between 2022 and 2047 will, at any given citation fraction, produce twice as many top-cited papers in absolute terms. Crucially, our estimations do *not* model the interaction between volume growth and citation quality: we have estimated the quantity and quality KPFs as separate specifications, without allowing volume growth to influence citation fractions or vice versa. This means Approach A, by holding volume constant, may generate an overly optimistic picture for G8 incumbents; one that could encourage complacency about the durability of quality leadership when volume positions are eroding.

Approach B addresses this by projecting the top-1% citation fraction from the quality KPF and then multiplying by the projected article volume from the quantity KPF to obtain quality-weighted output:

$$\dot{A}_{it}^{QW} = \dot{A}_{it} \times f_{it}^{top1}, \tag{18}$$

where $\dot{A}_{it}$ is projected article volume from Section 5.4.1 and f_{it}^{top1} is the projected citation fraction from equation (17). Each country's quality-weighted global share is its $\dot{A}_{it}^{QW}$ as a fraction of the world total.

Under Approach B, the sharp G8 attenuation reappears, driven not by the quality KPF

Table 16: Quality-Adjusted Projections: Direct (Approach A) and Quality-Weighted Volume (Approach B) Shares (%)

		Quantity share (%)			Approach A: top-1% share (%)			Approach B: QW share (%)		
ISO	Country	2022	2035	2047	2022	2035	2047	2022	2035	2047
CHN	China	22.57	29.52	27.67	21.18	22.72	23.84	24.46	32.93	28.50
USA	United States	14.01	6.85	3.00	21.50	19.23	17.10	19.11	8.10	2.75
GBR	United Kingdom	4.43	2.60	1.35	7.40	6.82	6.23	6.89	3.60	1.49
DEU	Germany	4.07	2.03	0.88	4.12	3.57	3.09	3.76	1.58	0.52
ITA	Italy	3.00	2.21	1.45	4.19	4.24	4.21	3.93	2.83	1.62
FRA	France	2.50	1.39	0.76	2.54	2.51	2.46	2.21	1.18	0.55
CAN	Canada	2.62	1.53	0.76	3.16	2.87	2.58	3.02	1.55	0.61
JPN	Japan	2.99	1.45	0.68	1.40	1.28	1.18	1.25	0.54	0.20
IND	India	5.48	8.34	9.33	3.14	3.55	3.93	3.90	6.49	7.08
BRA	Brazil	1.93	3.05	6.80	0.93	1.46	2.19	0.79	1.93	5.70
TUR	Türkiye	1.50	1.96	2.25	0.95	1.13	1.32	1.09	1.65	1.94
ZAF	South Africa	0.64	0.82	0.97	0.52	0.61	0.69	0.51	0.74	0.89
ISR	Israel	0.54	0.40	0.28	0.42	0.45	0.47	0.41	0.31	0.20
KOR	South Korea	2.22	1.78	1.30	1.50	1.58	1.63	1.43	1.17	0.78
MEX	Mexico	0.66	0.52	0.36	0.26	0.26	0.26	0.24	0.18	0.11

Notes: Quantity share = share of above-median sample output (77 countries); panel-common KPF elasticities $\hat{\sigma} = 2.30$, $\hat{\gamma} = -0.05$, $\hat{\phi} = 0.76$. **Approach A (primary)**: direct projection of top-1% citation share from PMG+CSD quality KPF ($\hat{\sigma}_q = 2.188$, $\hat{\gamma}_q = 0.039$, $\hat{\phi}_q = 0.015$) driven by factor input growth alone; citation baseline calibrated to 2020 (NSB S&E Indicators 2023; last year with sufficient citation accumulation), back-extrapolated to 2022 via quality KPF. **Approach B (qualifier)**: quality-weighted share = $\dot{A}_{it} \times f_{it}^{top1}$ normalised by world total, where $\dot{A}_{it}$ comes from the quantity projection and f_{it}^{top1} from the quality KPF; incorporates volume compounding dynamics. Input growth rates from Table 14.

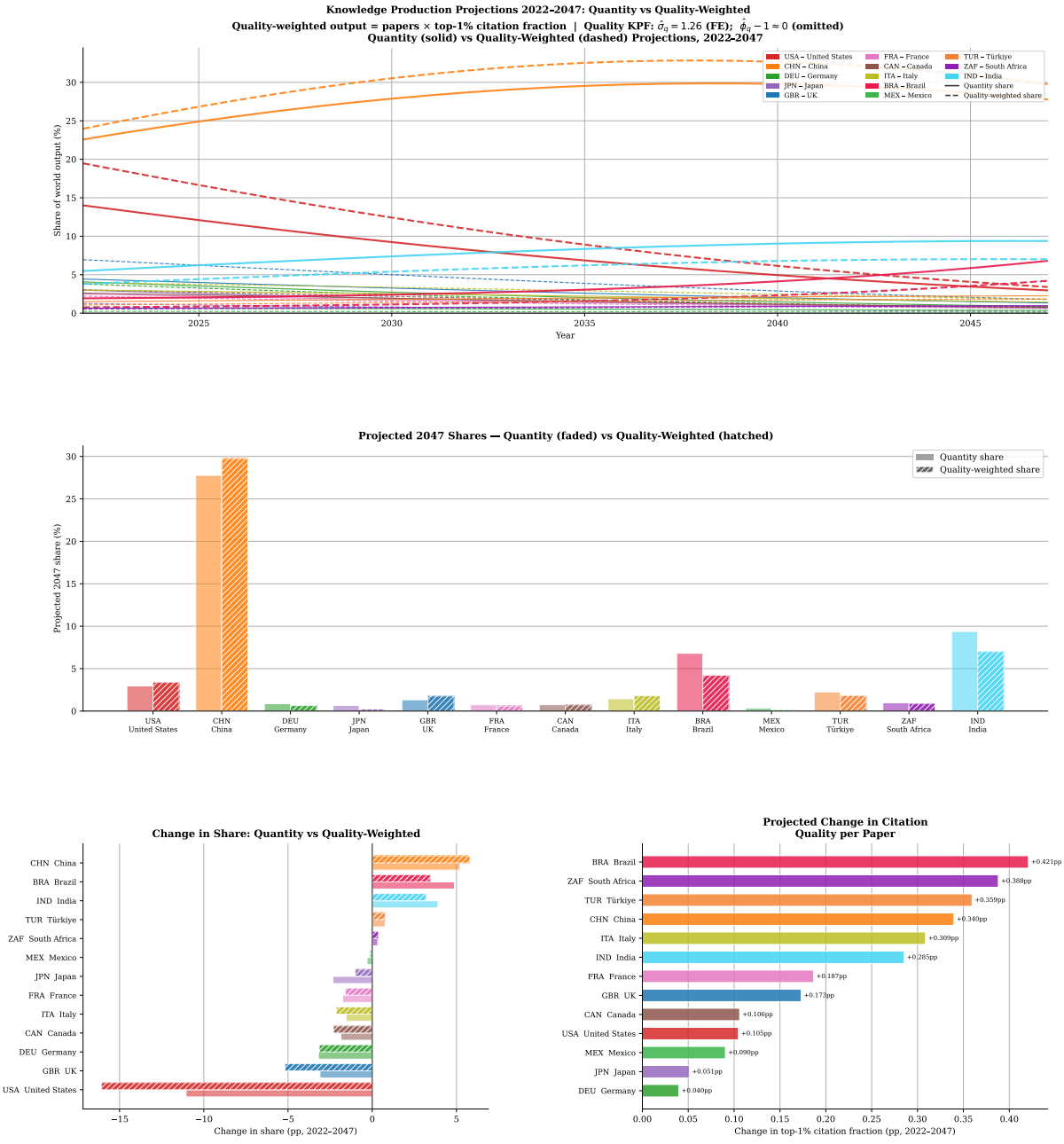


Figure 2: Quality-Adjusted Knowledge Production Projections, 2022–2047.
Notes: Quantity share (solid lines), Approach A top-1% direct share (dashed lines), and Approach B quality-weighted share (dotted lines) for selected focus economies. Approach A projects citation share from quality KPF factor inputs alone; Approach B additionally incorporates quantity KPF volume dynamics. See Table 16 and Section 5.4.2.

but by the volume compounding dynamics feeding through the multiplication. Table 16 (Approach B columns) shows the United States falling from 19.1% to 2.8% (−16.4 pp), the United Kingdom from 6.9% to 1.5% (−5.4 pp), and Germany from 3.8% to 0.5% (−3.2 pp). China's quality-weighted share rises from 24.5% to 28.5% (+4.0 pp). The difference between Approach B and Approach A (approximately 12 percentage points for the US, 4–5 pp for the UK) is precisely the contribution of volume-side compounding to the quality-weighted share calculation: it is the quantity compounding mechanism operating through the identity $\dot{A}^{QW} = \dot{A} \times f^{top1}$.

The policy lesson: security in quality leadership, but not complacency. The two approaches bracket the realistic range of outcomes and together deliver a coherent policy message. Approach A establishes that quality leadership, defined as share of world top-cited science, may be more durable than volume leadership: the US and G8 quality positions erode far more slowly than their volume positions when the projection is driven by quality KPF dynamics alone. This reflects the structural finding that citation quality depends on human capital, not knowledge-stock compounding, and human capital positions change slowly.

However, Approach A should not encourage complacency. The comparison with Approach B reveals that volume decline may transmit into quality-weighted share loss. While we have not modelled the interaction between volume and quality in the KPF, it is plausible that sustained volume decline will at some point erode quality leadership through mechanisms not captured here: reduced scale for experimental infrastructure, fewer graduate students, or thinning of the research workforce that generates the citation-quality premium. Approach B, by incorporating volume dynamics multiplicatively, may overstate the speed of this transmission, but it is likely the correct direction of effect. The gap between the two approaches (−4.4 pp vs −16.4 pp for the US) should be understood not as a range of uncertainty but as a decomposition: −4.4 pp is the pure quality-factor-input effect; the remaining −12 pp is what happens when we additionally account for the compounding decline in article volume that our estimation does not jointly model with quality.

6 Conclusion

This paper has estimated Knowledge Production Functions across 181 countries over 2003–2022 using heterogeneous-panel estimators that account for the cross-sectional dependence generated by a global science frontier. Five structural findings stand out.

First, the knowledge stock elasticity $\hat{\phi}$ is positive, large, and precisely estimated across all specifications and estimators ($\hat{\phi} \in [0.65, 0.76]$ in the preferred above-median sample). These estimates lie at the upper end of prior patent-based estimates in the liter-

ature ($\hat{\phi} \approx 0.3$–0.5 in levels specifications; Jones, 1995a) and establish that academic knowledge production exhibits stronger compounding than commercial innovation. The panel-average falls within the semi-endogenous regime ($\phi < 1$). The Pooled Mean Group estimator of Pesaran et al. (1999) formally confirms that the long-run relationship is homogeneous across countries (Hausman test: $p \approx 0.38$), validating the use of panel-mean elasticities and delivering a more precise estimate: $\hat{\phi} = 0.652$ (SE = 0.020) under PMG+CSD. A rest-of-world robustness check confirms that own-stock compounding is a genuine domestic effect: the coefficient is entirely stable ($\Delta\hat{\phi}_1 = -0.001$) when a rest-of-world knowledge stock spillover term is added under FE, with the spillover itself positive but imprecisely estimated. This cumulative advantage mechanism implies that the cross-country distribution of scientific output reflects accumulated investment histories that compound forward and are difficult to reverse.

Second, human capital quality exhibits substantial positive elasticity ($\hat{\sigma} \approx 2.30$ in the above-median sample under AMG, rising above 4.0 in the non-zero sample), suggesting markedly higher returns among countries at the extensive margin of science participation. The evidence is consistent with a two-stage model of science development: human capital quality drives initial entry into scientific activity, after which the compounding of accumulated knowledge becomes the primary engine of output growth.

Third, citation-quality output is governed by a fundamentally different production process than volume. Accumulated knowledge stocks are irrelevant to citation impact ($\hat{\phi}_q \approx 0.015$, negligible relative to $\hat{\phi} \approx 0.65$ for quantity), while human capital quality is the principal driver of citation impact ($\hat{\sigma}_q = 2.19$, $t = 11.2$, under the cross-sectional-dependence-corrected PMG estimator appropriate to this specification's I(0) dependent variable). A complementary country-level decomposition shows that time-invariant institutional characteristics also explain a substantial share of cross-country differences in citation quality, beyond what human capital predicts. Quantity and quality of scientific output are not simply different scales of the same activity; they are structurally distinct in their production economics, with quality additionally requiring institutional conditions, not only inputs, to be deployed effectively.

Fourth, the dominant finding from estimation is parameter homogeneity: the cross-country distribution of KPF elasticities is tight enough that the PMG pooling restriction is not rejected, and most pairwise country differences are not robustly distinguishable at $T = 20$. The homogeneity finding validates the panel-mean elasticities used in the forward projections and is itself an empirical result: cross-country differences in knowledge production shares are driven primarily by input accumulation rates and initial conditions, not by heterogeneous production technologies.

Fifth, forward projections to 2047 indicate a substantial redistribution of global knowledge production volume, consistent with the structural findings above. China's volume share rises from 22.6% to 27.8% (peaking near 29.8% around 2040), while the United

States falls from 14.0% to below 3.0%. India approaches 10% and Brazil 7%. Every non-China G8 economy loses volume share. On quality, the picture is more nuanced: projecting citation shares from quality factor inputs alone (Approach A), the US retains 17.1% of world top-cited science by 2047, far above its 3% quantity share. Only when volume dynamics are incorporated (Approach B) does the quality-weighted decline approach the severity of the volume decline. These projections are conditional on structural stability of the KPF and current input growth trajectories; they describe the consequences of the status quo, not an immutable forecast.

Several limitations temper these conclusions. The $T = 20$ estimation window limits the precision of country-specific coefficient estimates, and field-level disaggregation would allow identification of sector-specific KPFs and tests for directed technical change (Acemoglu, 2002). The patent stock proxy for product variety performs poorly in AMG and CCEMG specifications, a genuine identification challenge that calls for better empirical measures of the Jones (1995a) product variety concept in the context of knowledge production functions. The quality specification, while identifying the structural asymmetry with the quantity KPF, does not yet explore alternative quality measures such as field-normalised citation scores or quality-adjusted researcher counts. Finally, the NSB bibliometric data are sourced from Web of Science, which systematically under-covers non-English journals, regional publications, and the social sciences and humanities (Mongeon & Paul-Hus, 2016); this coverage bias may understate the true science output of India, Brazil, Türkiye, and South Africa relative to established Anglophone systems, and future work drawing on broader databases such as OpenAlex would provide a useful robustness check.

Online Appendix The Online Appendix contains five sections: (A) an extended treatment of the background literature; (B) the full set of country-specific KPF coefficient scatter plots for the 8 G8 and 7 emerging market economies (15 figures, 8 estimates each); (C) country-specific versus panel-mean projection comparisons, including full convergence-scenario results; (D) the methodological implications of rest-of-world knowledge stocks for the estimators employed, with full results; and (E) Westerlund (2007) panel cointegration test statistics.

Data and replication. All data are drawn from publicly available sources: NSB S&E Indicators 2023 (National Science Board, 2023); WIPO IP Statistics Data Centre (World Intellectual Property Organization, 2026); Penn World Tables 11.0 (Feenstra et al., 2015); World Development Indicators (World Bank, 2026). The data employed for the study (above-median sample: 1,799 observations, 90 countries, 2003–2022, 85 variables; non-zero sample: 3,479 observations, 181 countries, 2003–2022, 85 variables) and replication code is available.

Funding. No funding was received for this study.

Conflict of Interest. No conflict of interest affects this study.

References

Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *Journal of Economic History*, *46*(2), 385–406. https://doi.org/10.1017/S0022050700046209

Acemoglu, D. (2002). Directed technical change. *Review of Economic Studies*, *69*(4), 781–809. https://doi.org/10.1111/1467-937X.00226

Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2), 323–351. https://doi.org/10.2307/2951599

Aghion, P., Jones, B. F., & Jones, C. I. (2017). Artificial intelligence and economic growth. (23928).

Ben-David, J. (1971). *The scientist's role in society: A comparative study*. Prentice-Hall.

Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development: Evidence from aggregate cross-country data. *Journal of Monetary Economics*, *34*(2), 143–173. https://doi.org/10.1016/0304-3932(94)90047-7

Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, *110*(4), 1104–1144. https://doi.org/10.1257/aer.20180338

Bond, S., & Eberhardt, M. (2009). *Cross-section dependence in non-stationary panel models: A novel estimator* [Paper presented at the Nordic Econometrics Conference, Lund University, Sweden].

Borjas, G. J., & Doran, K. B. (2012). The collapse of the Soviet Union and the productivity of American mathematicians. *Quarterly Journal of Economics*, *127*(3), 1143–1203. https://doi.org/10.1093/qje/qjs015

Chudik, A., Mohaddes, K., Pesaran, M. H., & Raissi, M. (2016). Long-run effects in large heterogeneous panel data models with cross-sectionally correlated errors. In *Essays in honor of Aman Ullah* (pp. 85–135, Vol. 36). Emerald Group Publishing Limited. https://doi.org/10.1108/S0731-905320160000036013

Coe, D. T., & Helpman, E. (1995). International R&D spillovers. *European Economic Review*, *39*(5), 859–887. https://doi.org/10.1016/0014-2921(94)00100-E

Coe, D. T., Helpman, E., & Hoffmaister, A. W. (2009). International R&D spillovers and institutions. *European Economic Review*, *53*(7), 723–741. https://doi.org/10.1016/j.euroecorev.2009.02.005

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, *35*(1), 128–152. https://doi.org/10.2307/2393553

Ditzen, J. (2021). Estimating long-run effects and the exponent of cross-sectional dependence: An update to `xtdcce2`. *Stata Journal*, *21*(3), 687–707. https://doi.org/10.1177/1536867X211045560

Eberhardt, M. (2012). Estimating panel time-series models with heterogeneous slopes. *Stata Journal*, *12*(1), 61–71. https://doi.org/10.1177/1536867X1201200105

Eberhardt, M., & Bond, S. (2009). *Cross-section dependence in non-stationary panel models: A novel estimator* [MPRA Paper No. 17692, University Library of Munich].

Eberhardt, M., & Teal, F. (2011). Econometrics for grumblers: A new look at the literature on cross-country growth empirics. *Journal of Economic Surveys*, *25*(1), 109–155. https://doi.org/10.1111/j.1467-6419.2010.00624.x

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251–276. https://doi.org/10.2307/1913236

Fagerberg, J. (1994). Technology and international differences in growth rates. *Journal of Economic Literature*, *32*(3), 1147–1175.

Fedderke, J. W., & Dong, X. (2026). Brain drain and productivity growth: Evidence from South Africa, 1947–2019. *Economics of Transition and Institutional Change*, *00*, 1–27. https://doi.org/10.1111/ecot.70031

Fedderke, J. W., & Liu, Y. (2017). Schumpeterian and semi-endogenous productivity growth explanations. *Economics of Transition and Institutional Change*, *25*(1), 111–137. https://doi.org/10.1111/ecot.12114

Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015). The next generation of the penn world table. *American Economic Review*, *105*(10), 3150–3182. https://doi.org/10.1257/aer.20130954

Fu, X. (2015). *China's path to innovation*. Cambridge University Press. https://doi.org/10.1017/CBO9781107588462

Gerschenkron, A. (1962). *Economic backwardness in historical perspective: A book of essays*. Belknap Press of Harvard University Press.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature*, *28*(4), 1661–1707.

Griliches, Z., & Mairesse, J. (1998). Production functions: The search for identification (S. Strøm, Ed.), 169–203.

Guellec, D., & van Pottelsberghe de la Potterie, B. (2001). R&d and productivity growth: Panel data analysis of 16 OECD countries. *OECD Economic Studies*, *33*, 103–126. https://doi.org/10.1787/eco_studies-v2001-art12-en

Guellec, D., & van Pottelsberghe de la Potterie, B. (2004). From R&D to productivity growth: Do the institutional settings and the source of funds of R&D matter? *Oxford Bulletin of Economics and Statistics*, *66*(3), 353–378. https://doi.org/10.1111/j.1468-0084.2004.00083.x

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools. (8498).

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, *36*(1), 16–38.

Hall, B. H., Mairesse, J., & Mohnen, P. (2010). Measuring the returns to R&D (B. H. Hall & N. Rosenberg, Eds.). *Handbook of the Economics of Innovation*, *2*, 1033–1082. https://doi.org/10.1016/S0169-7218(10)02008-3

Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, *17*(4), 267–321. https://doi.org/10.1007/s10887-012-9081-x

Heinze, T., Jappe, A., & Pithan, D. (2019). From North American hegemony to global competition for scientific leadership? Insights from the Nobel population. *PLOS ONE*, *14*(4), e0213916. https://doi.org/10.1371/journal.pone.0213916

Hollingsworth, J. R., & Hollingsworth, E. J. (2011). Major discoveries, creativity, and the dynamics of science. In *Major discoveries, creativity, and the dynamics of science* (pp. 57–69). Remaprint.

Hu, A. G. Z., & Jefferson, G. H. (2009). A great wall of patents: What is behind China's recent patent explosion? *Journal of Development Economics*, *90*(1), 57–68. https://doi.org/10.1016/j.jdeveco.2008.11.004

Jones, C. I. (1995a). R&D-based models of economic growth. *Journal of Political Economy*, *103*(4), 759–784. https://doi.org/10.1086/262002

Jones, C. I. (1995b). Time series tests of endogenous growth models. *Quarterly Journal of Economics*, *110*(2), 495–525. https://doi.org/10.2307/2118448

Kapetanios, G., Pesaran, M. H., & Yamagata, T. (2011). Panels with non-stationary multifactor error structures. *Journal of Econometrics*, *160*(2), 326–348. https://doi.org/10.1016/j.jeconom.2010.10.001

Kortum, S. (1997). Research, patenting, and technological change. *Econometrica*, *65*(6), 1389–1419. https://doi.org/10.2307/2171741

Larue, J., Gingras, Y., Sugimoto, C. R., & Haustein, S. (2022). Bibliometric inequality during the covid-19 pandemic: Country-level publication patterns. *Quantitative Science Studies*, *3*(1), 1–21.

Leydesdorff, L., & Wagner, C. S. (2009). Is the united states losing ground in science? a global perspective on the world science system. *Scientometrics*, *78*(1), 23–36. https://doi.org/10.1007/s11192-008-1961-y

Madsen, J. B. (2008). Semi-endogenous versus Schumpeterian growth models: Testing the knowledge production function using international data. *Journal of Economic Growth*, *13*(1), 1–26. https://doi.org/10.1007/s10887-007-9023-x

Meng, Y., & Romanowich, M. (2021). Catching up or falling behind? China's changing science and technology trajectories. *Research Policy*, *50*(9), 104333. https://doi.org/10.1016/j.respol.2021.104333

Millot, V. (2009). *Trademarks as an indicator of product and marketing innovations* (tech. rep. No. 2009/06). OECD Science, Technology and Industry Working Papers. https://doi.org/10.1787/224428874418

Mokyr, J. (2002). *The gifts of Athena: Historical origins of the knowledge economy*. Princeton University Press.

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of web of science and Scopus: A comparative analysis. *Scientometrics*, *106*(1), 213–228. https://doi.org/10.1007/s11192-015-1765-5

Moser, P., Voena, A., & Waldinger, F. (2014). German Jewish émigrés and US invention. *American Economic Review*, *104*(10), 3222–3255. https://doi.org/10.1257/aer.104.10.3222

Narin, F., Carpenter, M. P., & Woolk, N. C. (1976). Scientometric building blocks toward a theory of science. *Scientometrics*, *1*(1), 37–44.

Narin, F., Hamilton, K. S., & Olivastro, D. (1997). The increasing linkage between U.S. technology and public science. *Research Policy*, *26*(3), 317–330. https://doi.org/10.1016/S0048-7333(97)00013-9

National Science Board. (2023). *Science and engineering indicators 2023* (tech. rep. No. NSB-2023-33). National Science Foundation. Alexandria, VA. https://ncses.nsf.gov/pubs/nsb202333

Nelson, R. R. (Ed.). (1993). *National innovation systems: A comparative analysis*. Oxford University Press.

Nelson, R. R., & Wright, G. (1992). The rise and fall of American technological leadership: The postwar era in historical perspective. *Journal of Economic Literature*, *30*(4), 1931–1964.

Pakes, A., & Griliches, Z. (1984). Patents and R&D at the firm level: A first look (Z. Griliches, Ed.), 55–72.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, *74*(4), 967–1012. https://doi.org/10.1111/j.1468-0262.2006.00692.x

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*(2), 265–312. https://doi.org/10.1002/jae.951

Pesaran, M. H., Shin, Y., & Smith, R. P. (1999). Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association*, *94*(446), 621–634.

Pritchett, L. (2001). Where has all the education gone? *World Bank Economic Review*, *15*(3), 367–391. https://doi.org/10.1093/wber/15.3.367

Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, *12*(2), 111–134. https://doi.org/10.1080/0964529042000239140

Quan, W., Chen, B., & Shu, F. (2017). Publish or impoverish: An investigation of the monetary reward system of science in china (1999–2016). *Aslib Journal of Information Management*, *69*(5), 486–502. https://doi.org/10.1108/AJIM-01-2017-0014

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, *94*(5), 1002–1037. https://doi.org/10.1086/261420

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5, Part 2), S71–S102. https://doi.org/10.1086/261725

Segerstrom, P. S. (1998). Endogenous growth without scale effects. *American Economic Review*, *88*(5), 1290–1310.

Tollefson, J. (2018). China declared world's largest producer of scientific articles. *Nature*, *553*(7689), 390. https://doi.org/10.1038/d41586-018-00927-4

Ülkü, H. (2004). R&d, innovation and economic growth: An empirical analysis [International Monetary Fund, Washington, D.C.]. *IMF Working Paper*, *WP/04/185*.

von der Heyden, M., Heinze, T., Galvão de Almeida, R., & Breznau, N. (2025). Decline of German and rise of North-American hegemony in science: Insights from Nobel prize nominations (Physics/Chemistry, 1901–1969). *PLOS ONE*, *20*(5), e0323103. https://doi.org/10.1371/journal.pone.0323103

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., Van Eck, N. J., Van Leeuwen, T. N., Van Raan, A. F. J., Visser, M. S., & Wouters, P. (2015). The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology*, *63*(12), 2419–2432.

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2012). The leiden ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology*, *63*(12), 2419–2432. https://doi.org/10.1002/asi.22708

Westerlund, J. (2007). Testing for error correction in panel data. *Oxford Bulletin of Economics and Statistics*, *69*(6), 709–748. https://doi.org/10.1111/j.1468-0084.2007.00477.x

World Bank. (2026). World development indicators. https://databank.worldbank.org/source/world-development-indicators

World Intellectual Property Organization. (2026). WIPO IP Statistics Data Center. https://www3.wipo.int/ipstats/

Young, A. (1998). Growth without scale effects. *Journal of Political Economy*, *106*(1), 41–63. https://doi.org/10.1086/250002

Yu, D., Shi, S., & Mccay-Peet, L. (2022). Examining the internationalization of Chinese scientific publications: Quantity and quality. *Journal of Informetrics*, *16*(3), 101307. https://doi.org/10.1016/j.joi.2022.101307

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- KPFAppendix.pdf
- KPFAppendix.tex